Ilkka Huhtakallio

Effects of wind on speech production, transmission and perception

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 13.1.2014

Thesis supervisor and instructor:

Prof. Ville Pulkki

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Ilkka Huhtakallio

Title: Effects of wind on speech production, transmission and perception

Date: 13.1.2014 Language: English Number of pages:10+78

Department of Signal Processing and Acoustics

Professorship: Acoustics and Audio Signal Processing Code: S-89

Supervisor and instructor: Prof. Ville Pulkki

Speech is the main communication method between humans. The properties of the environment and transmission path can have various effects on the quality of the signal transmission. The focus of this thesis is on the speech communication in the presence of wind with magnitude of around 12 m/s.
In this thesis a background theory of wind induced effects on sound are presented, and their relevance within the normal speech communication distances are discussed. In addition of the study on sound transmission, the theory of human speech communication mechanisms are presented in order to evaluate the wind induced effect on speech production, radiation as well as perception.
The main part of this thesis consist of an design and implementation of a large pendulum system inside an anechoic chamber. By means of the pendulum system a relative motion of the medium and attached source/receiver was attained. For the use as a source, a simplified artificial speech source was designed and built. For the use as a receiver, a mannequin with realistic ear molds and in ear microphones was used. The data acquisition system was designed and implement for multipurpose use. With this measurement system, several different measurements was conducted. The results showed that low-frequency tone did not have any effect on the production or radiation of higher tones with presence of the flow. The effect of flow on mouth radiation load was studied, but no clear result was obtained to indicate such a behaviour. The measurements showed, that pressure field variation inside the vocal tract are minor. A directivity pattern of the artificial speech source was then measured. The results shows that the flow has an effect of the radiation pattern. The radiation boosts in the upwind direction and attenuates at the most in the direction normal to the flow. Finally the most clear result was obtained with head and torso mannequin. The interaural spectral differences of the wind-induced noise at the entrance of the ear canals depends highly of the wind direction.

Keywords: Flow, wind, speech, communication, vocal tract, directivity, acoustics, measurements, pendulum, data aqcuisition

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Ilkka Huhtakallio

Työn nimi: Tuulen vaikutus puheen tuottoon, siirtoon ja havaitsemiseen

Päivämäärä: 13.1.2014 Kieli: Englanti Sivumäärä:10+78

Signaalinkäsittelyn ja akustiikan laitos

Professuuri: Akustiikka ja äänenkäsittelytekniikka Koodi: S-89

Valvoja ja ohjaaja: Prof. Ville Pulkki

Puhe on ihmisten tärkein kommunikointimenetelmä. Ympäristön ja siirtoväylän ominaisuudet voivat vaikuttaa eri tavoin puhesignaalin siirron laatuun. Tämän työn tarkoituksena on tutkia puhekommunikaatiota tuulessa, jonka voimakkuus on 12 m/s.

Tässä työssä on esitelty tuulen aiheuttamia ilmiöitä ääneen. Eri ilmiöiden merkittävyyttä arvioidaan suhteessa etäisyyksiin, joilla puhekommunikaation voidaan ajatella olevan mahdollista. Myös ihmisen puhekommunikaation mahdollistavat mekanismit on esitelty. Näiden valossa on käsitelty tuulen vaikutusta puheen tuottoon ja havaitsemieen.

Merkittävä osa työstä koostuu heiluriin pohjautuvasta mittausjärjestelmästä, jonka avulla voimakasta tuulta vastaava suhteellinen liike voitiin tuottaa kaiuttomissa olosuhteissa. Työssä on esitelty myös mittauksissa käytetyt keinopäät. Suuri osa mittausjärjestelmää koostui heilurin hallinta-, datan keruu- ja signaalinkäsittelyjärjestelmästä.

Useita mittauksia suoritettiin. Mittaustuloksista voidaan todeta, että virtauksella ei ole vaikutusta puhekaistalla olevien äänesten moduloitumiseen, kun matalataajuuksisen modulointiääneksen tasoa vaihdellaan. Virtauksen vaikutusta suun säteilykuormaan tutkittiin mittaamalla äänen käyttätymistä ääntöväylän sisällä. Äänikentän muutokset olivat vähäisiä, eikä suoraa yhteyttä säteilykuorman muutokseen voitu havaita. Yksinkertaistetun keinopään säteilyn suuntaavuutta mitattiin. Tulokset osoittivat, että virtaus vaikuttaa suuntaavuuteen. Säteily voimistuu vastatuulen suuntaan ja heikkenee voimakkaimmin kohtisuorasti tuulen suuntaan nähden. Merkittävin tulos mittauksista saatiin toisen keinopään avulla, jolla mitattiin virtauksen aiheuttamaa kohinaa korvakäytävän suulla. Tulokset osoittivat että taajuuspektri korvien välillä vaihtelee voimakkaasti tuulen suunnasta riippuen.

Avainsanat: Virtaus, tuuli, kommunikaatio, ääntöväylä, suuntaavuus, akustiikka, mittaukset, heiluri

Acknowledgements

This work is carried out in Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, during years 2012-2014.

I would like to take this opportunity to express my gratitude to Professor Ville Pulkki for his willingness to offer me the opportunity to excetute this research as a member of his research group. Thank you for supervising this work and for the valuable support, ideas, professional guidance and feedback on improving this thesis. I must admire your interest towards new research topics and will to spread the acoustical research in the lab beyond the main trends.

I would also like to thank the whole department staff and especially the spatial sound research group for warm and supportive atmosphere. The environment and equipments for the research at the lab are as good as one might think of, and the possibility to a widespread learning also from others research is impressive. I want to give special thanks to Javier Gomez Bolanos for many interesting discussions around acoustics and measurements, and Timo Lähivaara from University of Eastern Finland for sharing the common interest around the research topic.

A final thanks goes out to my family for all the love and support. I want thank my parents for giving me guidance of the importance of studying, and for all the support and understanding during these years. I want to show the deepest gratitudes to my two ladys; Katri - my better half - for the love, understanding and being there with me, and our daughter Saimi, for being the shine of the days and to remind me what really is important.

Otaniemi, 13.1.2014

Ilkka Huhtakallio

Contents

Abbreviations

PDA	Personal digital assistant
CFD	Computational fluid dynamics
SPL	Sound pressure level
DFT	Discrete Fourier transform
DADEC	Dummy head with adjustable ear canals
NI	National Instruments
SNR	Signal-to-noise ratio
DUT	Device under test

List of Figures

Chapter 1

Introduction

Speech is the primary communication method between humans, and we have a unique ability to use our voice in information transmission over acoustical channels as speech. The speech is transmitted trough an acoustical transmission path to another person, who receives the information by the use of hearing. Acoustic signal radiation and sound transmission in air are always at least partly involved in the normal speech communication regardless of the rest of the transmission path or technology involved. Whether the speaker, listener or both are outside and exposed to wind, disruptions due to an inhomogenous medium may arise in signal transmission.

It is known that the wind-induced noise on microphones can evoke significant quality drop on the performance of the microphone. Windshields are used with microphones to reduce the effect of the wind on microphones. Due to a size restrictions, the use of bulky windshields is not practical on small scale microphones that are used in mobile devices, hearing aids, compact video cameras and PDA devices. Lot of the wind related studies on acoustic field tries to tackle down, or at least understand better the wind-induced noise problems and how should the microphones and their housings be designed to minimise performance vulnerability in windy conditions. Another important field studying the effect of wind on sound is atmospheric acoustics. The wind speed gradients are used to estimate the sound propagation in atmosphere, where the sound refraction can introduce major effect on sound distribution into an environment on a larger scale.

Not much attention has been paid into the effect of a moving medium on sound radiation and transmission on smaller scale, that is the case in speech communication outdoors with high wind present. Respectively the wind-induced noise perceived by human hearing mechanism and the disturbances of such a noise on perception of other sounds have not been under a wide research. The hypothesis behind this work is that a high wind may introduce effect on sound radiation and propagation, and there might be phenomena that hasn't been studied in the context of speech communication.

Due to a nature of the wind, or in the case of more controlled airflows, the reproduction of them in laboratory environments is not straightforward with simple manners. The flow generation task is not simplified at all, if it needs to be done quietly in acoustically controlled space.

The aim of this thesis is to introduce and discuss the phenomena related to wind-induced effects on speech communication, and first of all to design and build a flow generation method in anechoic conditions. The background theory gives the basic knowledge of the individual phenomena of the hypothesis, and with the results obtained from the measurement, the significance of these phenomena forms the basic for the discussion.

This thesis is structured in two main categories; the theory and the measurements.The former is browsed trough in chapters 2-3, after which the chapters 4-6 covers the description of the measurement procedure and the results. Finally, the chapters 7-8 are left for a discussion and conclusions. Chapter 2 discusses the nature of the sound, atmospheric acoustics and fundamentals of aerodynamics. The aim of these topics is to introduce the basics of the phenomenas that may affect the acoustics of a problem with wind present. Chapter 3 covers the speech production and hearing mechanism with some details in addition with introduction to masking effects of broadband noise. Chapter 4 summarises the methods that have been used in flow generation in others studies related to wind-induced noise. Chapter 5 describes the measurement system used in this work. Chapter 5 includes introduction of pendulum physics and hardware. Chapter 6 finally presents the measurements conducted with the measurement system and the results of the measurements.

Chapter 2

Fundamentals of acoustics and aerodynamics

Sound is an acoustic wave with frequency in the range of hearing. Acoustic wave is longitudinal wave that propagates in elastic mediums, like gases and liquids. Usually the sound that surround us is a superposition of waves arriving from different sources. The propagation of the sound wave realizes as an alteration of pressure and particle movement of the medium. As such alterations reach our ears, they produce a sensation in our hearing mechanism, that is also informally called a sound.

2.1 Characteristics of sound

The sound can be described by its characteristics in different ways. A first categorization is to split the sound in wanted and unwanted sound. Unwanted sound is called a noise, that can disrupt the hearing of the wanted sound.

Another categorization is to describe the sound wave by its temporal nature to periodic and aperiodic sound waves. In periodic wave, the successive disturbances of the medium occurs at regular intervals in a repetitive pattern of cycles, that share the same shape. The time duration of one cycle is fundamental period T and the numbers of cycles occurred in one second is fundamental frequency f, so that $T = 1/f$. A glottis signal that will be described in chapter 3 is an example of periodic sound. In contrast to periodic sound is an aperiodic sound, where such periodically patterns does not exist, and successive disturbances occurs in random manner. An example of aperiodic sound is an wind-induced noise that will be introduced later in this chapter.

Closely related to periodic properties of sound, it can be described by its spectral content, which relates how the sound energy is spread over different frequencies. In case of periodic signals, the energy is spread over a harmonic structure, where frequencies are some integer multiples of fundamental frequency f_0. In aperiodic

signal, sound energy is spread more equally over a wider frequency band instead of single frequency components. Couple of other important discriminations are level and the spectral balance of the sound.

The normal sound scape that surround humans is a complex mixture of sound waves, that consist of wanted sound and noise, who both can be either periodic or aperiodic and have different duration, spectral balance and overall level.

2.2 Sound radiation and propagation

The sound is usually produced by some vibrating object which is in contact with the air. As the vibrating object is moving in one direction, it pushes air in front of it and compresses it. Likewise when the object moves to an opposite direction, the surrounding air expands to fill up the space left by the retreating object. Thus a series of compressions and rarefactions of the air are produced related to motion of the object. Due to an elasticity of the air, compression and rarefaction of the air near the object doesn't remain stationary, but move away from the object. The compression and rarefaction of the air produces a time and location dependent pressure variations above and below the average atmospheric pressure p_0. The instantaneous sound pressure p, or excess pressure, at a certain point is the total instantaneous pressure minus the atmospheric pressure, $p = p_{\text{tot}} - p_0$. The effective sound pressure p_{rms} at certain location is root-mean-square (rms) value of the excess pressure over some time interval. The term "sound pressure" is commonly used instead of effective sound pressure. The sound pressure p is usually presented as sound pressure level (SPL), L_{p} in decibels, $L_{\text{p}} = 20\log_{10}(p/p_{\text{ref}})$ dB, where p_{ref} equals 20 μPa.(24)

In general the sound propagation takes place in three dimensions and is governed by acoustical wave equation. The general three-dimensional acoustical wave equation in cartesian coordinates is

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{2.1}$$

This equation relates the second derivative of excess pressure p with coordinates x,y and z and second derivative on excess pressure with time t trough the square of the speed of sound c (12). In the derivation of 2.1 the following assumptions are made(12; 24) :

1. The amount of matter entering the borders of small volume is equal of the increase of the matter inside the volume, so no new matter is created nor the matter is destroyed in the process.
2. The net force applied to the small volume will get the matter inside the volume into a acceleration in direction of the net force. The net force arises from pressure differences on the sides of the small volume, the pressure gradient.

3. Although the pressure and density variations of a small volume introduces a change in temperature inside the volume, the variations in sound waves are so rapid that heat energy is not transferred in or out of the element, so the process is adiabatic.

4. The medium is not moving.

The speed of sound c for a ideal gas is given by $c = (\gamma RT)^{1/2}$, where γ is ratio of specific heats of the gas, R is the gas constant and T is absolute temperature. The speed of sound in air near normal room temperature is approximately $c = c_0 + 0.6T_c$, where $c_0 = 331.6$ m/s and T_c is temperature in degrees Celsius. The acoustical wave equation given in equation 2.1 has the same form for particle displacement ξ, particle velocity u, fluctuating density ρ, or the fluctuating temperature T (12). The wave equation can be transferred into different coordinate systems depending on the problem geometry.

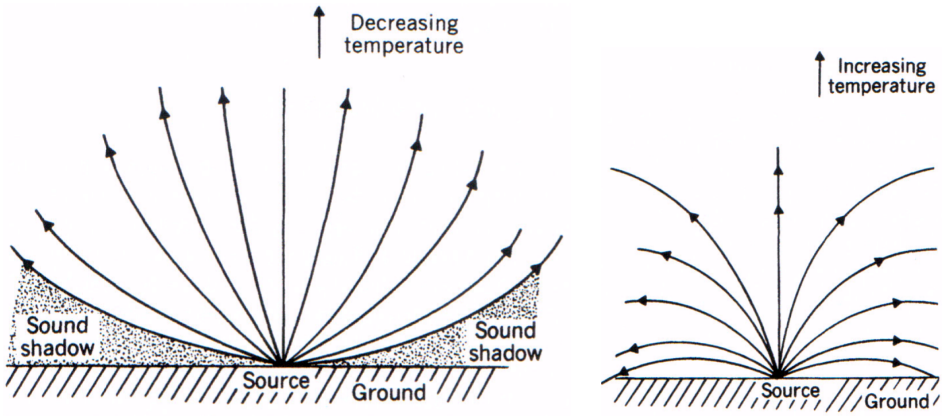
The sound propagation from a source to receiver is determined by the properties of the medium and all the obstacles and surfaces in between and in vicinity of the transmission path between the two points. If a simple source with its dimensions small compared to wavelength is vibrating, a spherically spreading sound wave is produced and it is propagating in all directions. The sound power in such a wave is distributed within the larger area as observation point is moved further away from the source, so the sound intensity is decreased and the sound is attenuated.

The propagation of the sound can be described by ray trajectories, as it is done in ray acoustics. Ray acoustics is also called the geometric acoustics and it is probably the easiest and best way of thinking about sound propagation problems (12). The ray trajectories presents the direction of the propagating wave, such that the trajectory is perpendicular to the wave fronts which are described by the constant phase regions of the wave. The direction of average energy flux follows that of the ray trajectories, and the amplitude of the field can be obtained from the density of the rays. The overall transfer function of the sound between the transmission and receiving points is a sum of all the rays from different paths reaching the receiver. The rays from the different paths arrives the receiving point differently attenuated at the different times. The time difference of the rays turns up as a frequency dependent phase difference of the sound waves. Depending on the phase difference either a constructive or a destructive interface occurs.

2.3 Sound propagation in moving medium

The sound propagation in moving medium has many complex features, and detailed analysis of most of them are out of the scope of this thesis. Some common phenomenas are introduced, when relative velocities of the sound sources, receivers and propagation medium are different.

The ray acoustics is often used to describe the phenomenas related to the outdoor



(a) Normal temperature lapse. (b) Temperature inversion.

Figure 2.1: Refraction due to temperature gradients. Figure adapted from (12)

sound propagation. The outdoor sound propagation, or atmospheric sound propagation, is of special interest in environmental acoustics which is concerned with the control of sound in outdoor environment. It includes geometric spreading and atmospheric absorption, propagation over ground and water, reflection and diffraction by obstacles and refraction due to non-homogenous medium. Refraction takes place when the speed of sound in changes function of the altitude. That happens because of the temperature and wind gradients above the ground. The sound refraction due to the moving medium is most commonly described effect of the wind on sound propagation, and it is usually studied as a part of the atmospheric sound propagation.

The refraction effect can have various strengths on either upward or downward direction. Upward refraction takes place in normal temperature lapse as shown in figure 2.1a and downward refraction takes place in a case of temperature inversion (figure 2.1b). In case of a wind, the mean wind velocity increases with height above the ground. In the direction of the wind, speed of sound increases with height and sound if refracted downwards. Respectively the speed of sound with height decreases in direction against the wind and sound is refracted upwards. The sound refraction due to wind in idealistic case is shown in figure 2.2. The most convenient profile for physical interpretation and mathematical computation is one were the sound velocity changes linearly with height above the ground (12):

$$c(h) = c_0(1 + \xi h) \tag{2.2}$$

where h is height and ξ is the normalized sound velocity gradient $(dc/dh)/c_0$. The negative sound gradients means upward refraction and creation of shadow regions, where no direct sound is transmitted. Because of the atmospheric turbulence and diffraction, the acoustic shadow regions are never complete, as an optical one would be (28). The presence of shadow region means that the sound level decreases faster that would happen in case of distance attenuation only. According to (28) when a combination of a slightly negative temperature gradient, strong upwind propagation

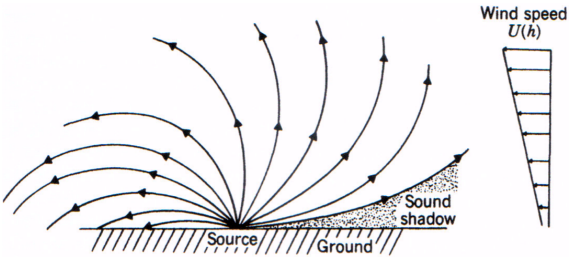


Figure 2.2: Refraction of sound due to increasing wind speed with altitude. Figure adapted from (12)

and air absorption, the reduced sound level up to 20 dB more than expected from spherical spreading was observed in carefully monitored experiments at 640 m away from 6 m-high source over relatively hard ground. This experiment shows that attenuation in shadow zones can be significant, but are most likely to happen far beyond the normal speech communication distance. The distance of the boundaries of the shadow zones can be located by ray tracing and geometrical considerations. Such analytical inspection gives valuable information of the scale of the shadow zone effect. As shown in figure 2.3, the distance of the shadow zone boundary is around 300 m for normalized sound speed gradient of 0.0001 m^{-1} if the source and receiver are 1 m above the ground. In presence of the high temperature and wind gradient, shadow zones can occur as close as 100 m (22) in the upwind direction. In the direction of the downwind, the wind will counteract the effect of temperature lapse and the shadow zone is destroyed.

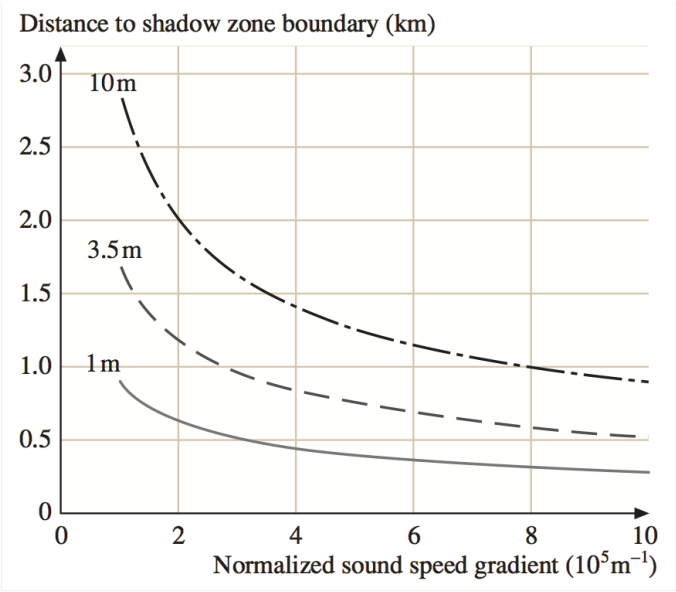


Figure 2.3: Distances to shadow zone boundaries for linear speed of sound gradient for equal source and receiver heights: 1 m (solid line), 3.5 m (broken line) and 10m (dot-dash line).Figure adapted from (28)

2.4 Sound in irregular flow

As summarized by Ingård (22), most of the short-range propagation effects are not related to sound speed gradients, but instead of the irregularities in the wind structure. The linear sound speed gradients rarely exist in windy conditions. The characterization of atmosphere by vertical profile in presence of the wind, are usually an average value over a longer time period, say around 10 minutes or more. The real atmosphere is in constant movement, and its profile changes continuously. Wind is generally fluctuating in speed and direction. Wind speed fluctuations comes up as gusts with different degree. Gustiness can be described by the gust factor $G = V_g/V_{avg}$, that is the relation of the velocity of the gusts V_g to mean wind velocity V_{avg}. The gustiness is caused by the atmospheric eddies of various sizes, that are developed as a wind flows over rough terrain, buildings, mountains or other irregularities. The fluctuating eddies can occur in a scale of several hours to a much faster scale of minutes or seconds. The faster scale fluctuations are known as a atmospheric turbulence. The subject of turbulent flow in general is deep, extensively studied, but still imprecise (3). In turbulent flow, fluid particles moves often in loops corresponding to swirls or eddies with arbitrary oriented axes. The scale of the turbulent eddies expands from millimeters up to 1-2 km. The kinetic energy of the large scale eddies is continuously transferred to smaller ones, as long the energy is completely dissipated by viscosity.

In the review work by Ingård (22) it was shown that a gusty turbulent atmosphere can attenuate the sound propagation along the 100 m distance from 6 dB at 250 Hz up to 12 dB at 4 kHz. He explained the phenomenon that a sound beam striking an eddy will be deviated in one and the same direction no matter which side the sound ray passes the axis of the eddy. The deviation of the ray may be of order 2 °or even more, and passing of multiple consecutive eddies may very well have notable effect. Another effect is due eddies rotating in different directions. They will cause deviations in different directions, and summation of different propagation paths may introduce destructive interference, at least momentary (22). Other literature source(28) explains that the importance of the eddy scale on sound propagation can be estimated by Bragg diffraction effect. It relates the spatial periodicity D (figure 2.4) to the wavelength λ and angle of scattering θ, so that $\lambda = 2Dsin(\theta/2)$.

2.5 Doppler effect

The familiar effect on sound due to a relative motion of the sound source and observer is the changing of the pitch of the observed sound. The effect is called a Doppler effect. One common example of the Doppler effect is an emergency vehicle with a siren passing by the receiver with high velocity. The effect is the same whether the sound in a medium is received by an observer in motion with respect to the medium.

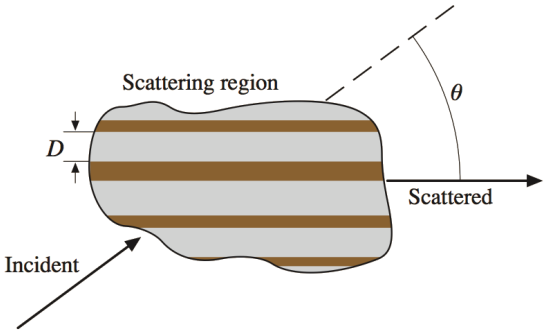


Figure 2.4: Scattering from turbulence. Figure adapted from (28)

The frequency f_o at the observation point is

$$f_o = \frac{c + V - v_o}{c + V - v_s} f_s \tag{2.3}$$

where c is the velocity of sound in the medium, V is the velocity of the medium in the direction in which the sound is traveling, v_o is the velocity of the observer related to the medium, v_s is the velocity of the source related to the medium and f_s is the frequency of the source (24). The case of the constant pitch change would occur only if the direction of the relative motion would follow the line connecting the source and the observer. In such case the observed pitch would be stationary upshifted until the source passes the observer, after which the observed pitch would jump immediately down and stay stationary as the source moves away. This is not the case when the source and observer are passing each other by some distance. In such case the radial velocity component does not remain constant but varies with the function of the angle between the line of sight and direction of velocity. Astronomer John Dobson explained the phenomena by saying: "The reason the siren slides is because it doesn't hit you" (27).

The more general formulation of the Doppler effect can to be done in time domain. It is needed in the case of non constant velocities. If the propagation and distance attenuations are not taken account, so that the pressure magnitude at the source $P_s(t_s)$ will equal the pressure magnitude at the observer $P_o(t_o)$, the distortion in signal due to a relative motion is then due the time warping, whereby the source and observer magnitudes are distributed on different time scales. The time domain mapping function can be solved from the problem geometry that satisfies the function $t_s = G(t_o)$, where t_s and t_o are time instances of the wave emission at the source and reception of the wave at the observer location. The wave as seen by the observer will be then given by (9)

$$P_o(t_o) = P_s(G(t_o)). \tag{2.4}$$

2.6 Basic principles of aerodynamics

The interaction between motion of the air and solid bodies are studied in the complex field of aerodynamics. Many of the aerodynamical studies and books are focused to explain why airplanes fly and what phenomenas are related in flying. The basic principles of aerodynamics cannot be neglected, or at least their existence is necessary to know, when a effects are search in an interaction of the moving medium and sound in presence of radiators or receivers with solid bodies. In this thesis, the aerodynamical phenomena can be used to describe the behavior of the air motion near the human heads to explain the possible effects of the wind on speech communication. Next the two fundamental principles of aerodynamics, Bernoulli's principle and law of continuity, are introduced after Anderson (3) and Smith (30).

Bernoulli's principle is derived from the conservation of energy, and it states that increase in velocity of the flow increases its kinetic energy and dynamic pressure and respectively, to equate the total amount of energy in the system, decreases the static pressure and potential energy. The principle that is also commonly noted as a Bernoulli's equation, can be expressed mathematically by:

$$p_{\mathrm{s}} + q_{\mathrm{d}} = p_{\mathrm{tot}} = \mathrm{const} \tag{2.5}$$

where p_{tot} is the total pressure, p_{s} is the static pressure and q_{d} is the dynamical pressure. The dynamic pressure is defined by $p_{\mathrm{d}} = \frac{1}{2}\rho V^2$, that is actually the kinetic energy per unit volume. The other important principle, law of continuity is really the law of conservation of matter applied to a moving fluid. The law states that the rate at which matter enters the system much equal the rate at which mass leaves the system. This can be written in the relation

$$\rho A V = \mathrm{const} \tag{2.6}$$

where A is the area and V is the velocity of the flow. If the velocity of the flow is low compared to the speed of sound, the density ρ of the air doesn't change, and the flow is incompressible. Now if a solid object is put into a free flow with velocity V_∞, the local area of the flow reduces because the presence of the object, and thus the local velocity of the flow must increase for equation 2.6 to hold. Because of the increased velocity, the kinetic energy and dynamic pressure is increased that leads the local decrease of the pressure according to Bernoulli's equation 2.5. If the object in the flow is symmetrically related to the flow, pressure drop are equal on both sides of the object and no force is applied to the body normal to the flow direction.

In opposite to basic problems of aerodynamic, in our study it is not that important to pay attention on forces created on bodies. It is more important to be able to predict the local pressure and flow velocity maximum points around the head with different wind angles. Basic equations to for these predictions are given in following section.

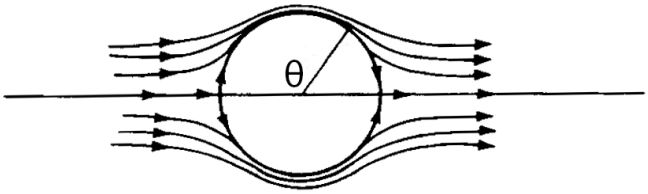


Figure 2.5: Schematic of the flow around sphere.

2.7 Nature of flow around objects

The pressure and velocity distribution around three dimensional object can be described by simplified example of flow around the sphere. The tangential velocity V_θ on the surface of a sphere is given by (3, p.426)

$$V_\theta = \frac{3}{2} V_\infty sin\theta \tag{2.7}$$

where V_∞ is a free stream velocity of the flow and θ is angle of the point related to the upwind direction as shown in figure 2.5. The $\theta = 0$ point is called a stagnation point, where the local velocity over the surface is zero. From this equation it can be seen that maximum velocity occurs on both sides of the sphere, when $\theta = \pm 90$. The pressure coefficient C_p is a dimensionless quantity that describes how much the local pressure p differs from the free stream pressure p_∞ in multiples of dynamic pressure q_d , namely $p = p_\infty + q_\mathrm{d} C_\mathrm{p}$. The pressure coefficient around the sphere is given by equation (after Anderson (3, p.426))

$$C_\mathrm{p} = 1 - \frac{9}{4} sin^2\theta \tag{2.8}$$

This equation states that the pressure around the sphere varies from maximum of one to minimum of -1.25 times the dynamic pressure of the flow. The maximum pressure occurs at the stagnation points and minimum pressures at the both sides of the sphere, where is the location of the maximum velocity. To get a view of the magnitude of the pressure variation in this research with the presence of the 12 m/s flow, the dynamic pressure is $q_d = 88.56$ Pa, that equals approximately the pressure of one cm H_2O.

Chapter 3

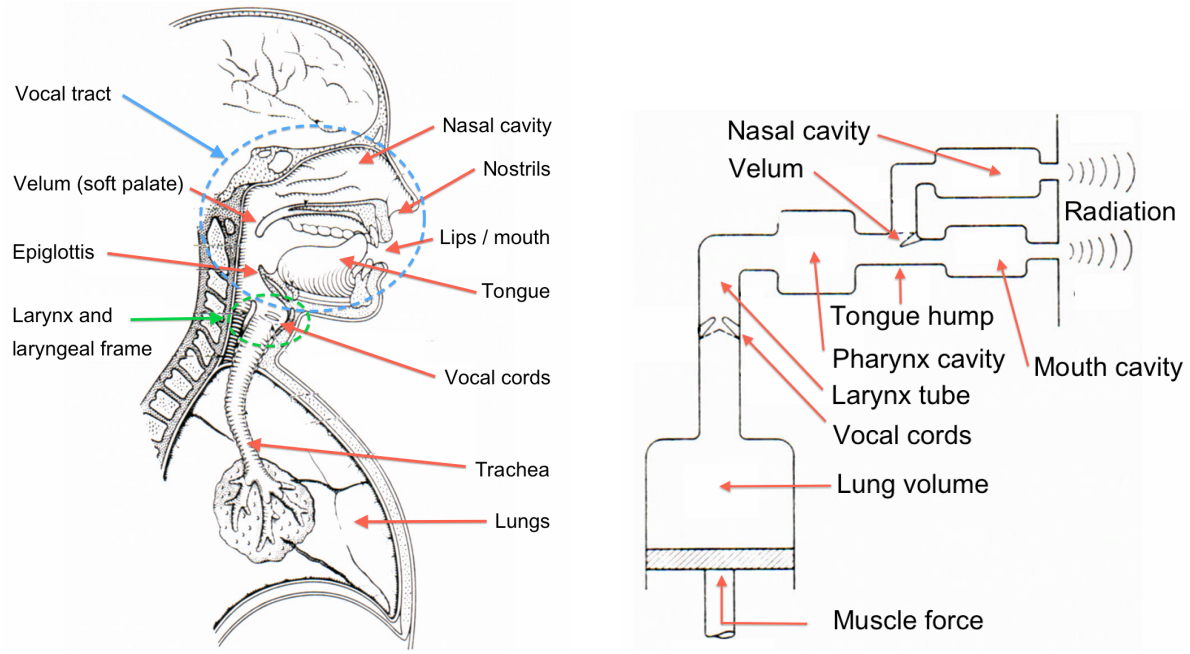
Speech communication

Speech is the primary communication method within humans, and we have a unique ability to use our voice in information transmission. The speech information is transmitted trough some transmission medium to another person, who receives the information by the use of hearing. Acoustic signal radiation and sound transmission in air are always at least partly involved in the normal speech communication regardless of the rest of the transmission path.

The speech information to be transmitted is first contained in neural signals. These neural signals control our muscles that are involved in speech production. Same muscles are involved in breathing and other physiological task in full synchronization with speech production. The acoustic feedback of the hearing mechanism and the kinesthetic feedback of the speech musculatures is used to develop, control and maintain the motorized behavior of speech. The speech is a formed from a set of words that are created out of the phonetic combination of a limited set of vowel and consonant speech sound units. The speech sounds can be categorized in a voiced and unvoiced sounds, which refer to an articulation process of fundamental sound production mechanism. In case of voiced sounds the sound is produced with a assistance of vocal cord vibration. Voiceless sounds are produced by the turbulence from the local constriction along the vocal tract, or instant release of the pressure build ups made by totally closed constrictions of the vocal tract. This study focuses on the voiced sounds. (15)

3.1 Human speech production mechanism

Speech production mechanism can be divided in three main subsystems, 1) lungs and trachea, 2) larynx and 3) vocal tract, as shown in figure 3.1a. These three subsystems forms a highly controllable mechanism that is used to produce sounds needed for language. The lungs acts as an air compressor, and the trachea is an air canal to a larynx, where an air flow is modulated by the vocal chords in case of a voiced speech. Modulation of the airflow produces a pressure variations creating a train of glottis



(a) Human speech production mechanism

(b) Simplified schematic of functional components of speech production mechanism

Figure 3.1: Overview of human speech production mechanism shows lungs and trachea, larynx together with glottis and vocal tract. Figures adapted from (15)

pulses. So the sound of a voiced speech sounds are produced in larynx. Vocal tract filters the glottis pulse train to its final shape, and it is radiated to a surrounding air trough mouth and nostrils. In this section these three subsystems and acoustic phonetics are browsed trough with some details.

3.1.1 Lungs and trachea as a power supply

Lungs are passive structures, and they are not able to inhale or exhale by themselves. Lungs are a pair of spongy tissue offering a large area to absorb oxygen from air to bloodstream. Lungs are located on either side of the chest and they are covered by a thin tissue layer called the pleura. The pleura is enclosed on the sides by a rig cage and from the bottom by a diaphragm. Diaphragm is a dome-shaped muscle connected to the bottom of the rib cage. Inhalation is accomplished by expanding the volume of the pleura by contraction of the diaphragm so that it lowers and flattens out. The volume may be increased even more by expanding the rib cage. The increased volume of the lungs causes a fall in pressure and air starts to flow into the lungs via nostrils, nasal cavity, velum port, and trachea. The capacity of the lungs of an adults are 4 to 5 liters. By relaxation of the diaphragm, elastic lungs starts to suppress and air pressure increases in the lungs.

The trachea, also known as a windpipe, is a rigid construction joining the lungs to the larynx. At the bottom the trachea branches into a left and right bronchus

that lead in to the lungs. At the top the trachea joins the larynx. The trachea is made of ring shaped cartilages that are joined together by tissue. This structure is approximately 2 cm across and 12 cm long tube, that can bend along with the head movements. Because of the relative large cross sectional area of the bronchial and tracheal tubes, the pressure drop across them is small. Therefore the lung pressure and subglottal pressure at the larynx are nearly the same. (15; 25)

The breathing cycle is used as a source of energy for speech production. Speech is produced during the exhale phase. The lung pressure is maintained steady by slow contraction of the rib cage and relaxation of diaphragm. The loudness of the speech is controlled by the lung pressure. Production of the vowel sounds at the softest possible level requires lung pressure of the order 4 cm H_2O [1]. Very loud high pitched sound may require up to 20 cm H_2O lung pressure. (15).

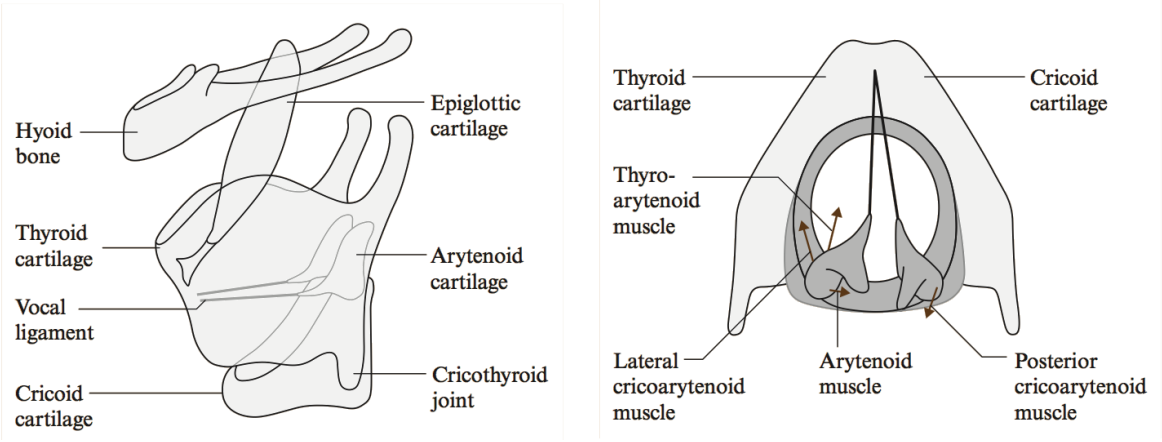
3.1.2 Source of voiced sound

The voiced sounds of speech are produced by vibrational behavior of vocal cords. Vocal cords are composed of two lips of ligament and muscle that are housed by a cartilaginous frame. This structure is stretched horizontally across the larynx at the top of the trachea. Tension as well as position of the back end of vocal cords are controlled by the set of muscles. By moving the ends of the vocal cords together or apart one can open or close the free air path to the lungs. The slot-like orifice between the vocal cords is called glottis. The glottis separates two air cavities on either side of the glottis. subglottal cavity is below the glottis at the top end of the trachea and supraglottal cavity is right above the glottis.

During the breathing the ends of the vocal cords are apart and open glottis lets the air move back and forth freely. If the ends of the vocal cords are pulled together, an air tight sealing is produced by the closed glottis. By closing the glottis a pressure can be built up in the thorax by increasing the lung pressure by means of the diaphragm and intercostal muscles. The pressure build up at the subglottal cavity acts as a energy container for the voiced speech production. By controlling the subglottal pressure and the tension of the vocal cords, air starts to flow trough a glottis. The forced airflow evokes the vibration of the vocal cords. The vibration modulates the airflow producing a buzzing sound, that is the excitation for voiced speech. Production of sound in this manner is called phonation (15).

The vocal cord oscillatory motion is result of repetitive closing and opening of the glottis, that repeats four phases within a cycle; the closed phase, opening phase, open phase and closing phase (6). The vibrational behavior is determined by the tension and mass of the vocal cords, the width of the glottis, a pressure difference across the glottis and volume velocity of the flow trough it. In closed phase the

[1] The cm $H_2O$is a unit of pressure used in speech sciences. It is defined as the pressure exerted by a column of water of 1 cm in height at 4 °C at the standard acceleration of gravity. One cm H_2O equals 98.0665 pascals.

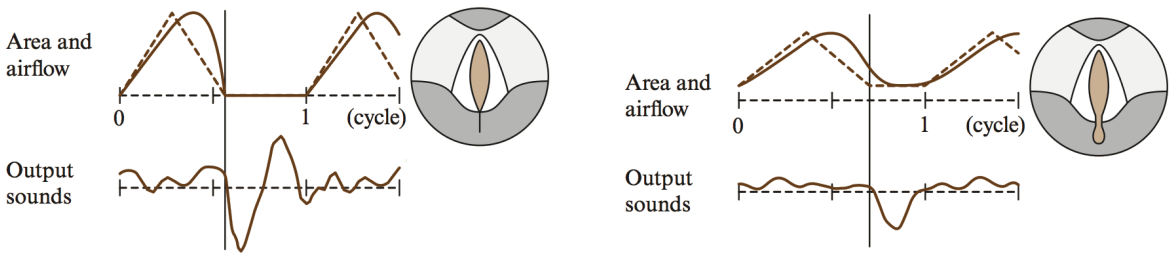


(a) Laryngeal framework (b) Glottal action

Figure 3.2: a) a view of the frame supporting vocal cords. b) top view including muscle action for abduction and adduction of the glottis. Figure adapted from (6)

subglottal pressure is increased. When the transglottal pressure difference exceeds a certain level it forces vocal cords apart with a lateral acceleration. In opening phase the velocity of airflow starts to increase and the glottis opens more and releases more air trough it. In open phase the glottal volume flow reaches its maximum value. According to the Bernoulli relation and developed volume velocity, local pressure starts to reduce in the orifice producing a force between two cords dragging them to a proximate position. The closing phase is usually shorter than a opening phase, but it may vary widely individually and related to a intonation and intensity of speech. As glottis is totally closed, the pressure starts to build up again in the subglottal cavity and a oscillation cycle starts from the beginning. The period of the oscillation is normally shorter than the natural period determined by the mass and the stiffness of the cords. The cords are driven in a forced oscillation producing a quasi-periodic pulses of air passing from the variable area orifice to excite the acoustic system above the glottis. (15; 6)

The relation of the glottal area, volumetric airflow, a glottal excitation sound and glottal shapes at peak openings are shown in figure 3.3 in case of modal phonation and breathy phonation. In modal phonation a total closure of the glottis within a cycle takes place, while in breathy phonation the glottis never closes totally within a cycle and constant airflow trough glottis shows as a DC component in the glottal waveform. In modal phonation, the opening and closing phase of the vocal cords shows a constant rate of change in area, while the airflow shows a slow rising rate followed by a transition to a faster decay rate and a instant discontinuous decrease to zero as glottis closes. A quick interruption of the airflow is the main source of vocal tract excitation, as shown in the figure 3.3a (6). In breathy phonation the changes in airflow are smoother and produced excitation sound is weaker. The voice quality can be altered by the means of control of the vocal cord oscillation. The pitch and spectral content of the glottis excitation changes as a ratio of the open phase to



(a) Modal phonation (b) Breathy phonation

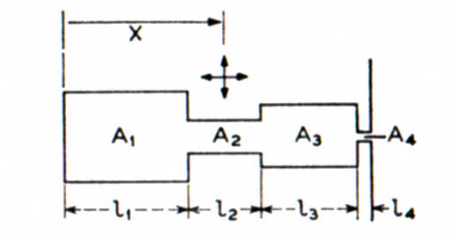
Figure 3.3: Relation of the glottal area (dashed line), airflow and sound produced during 1.5 glottal cycles. The shape of the glottal openings are shown in circles. Figure adapted from (6)

closed phase within a cycle changes, or the ratio of the closing rate to a opening rate changes. For normal voice effort and pitch the ratio of the open time to total period (duty cycle) is commonly of the order of 0.3 to 0.7. The glottal waveform of normal phonation has relatively rich spectrum with roughly 12 dB/oct attenuation in the higher frequencies. In neutral speech there are some cycle-to-cycle variation in the oscillation that can be observed as a frequency and amplitude permutations. (15; 6)

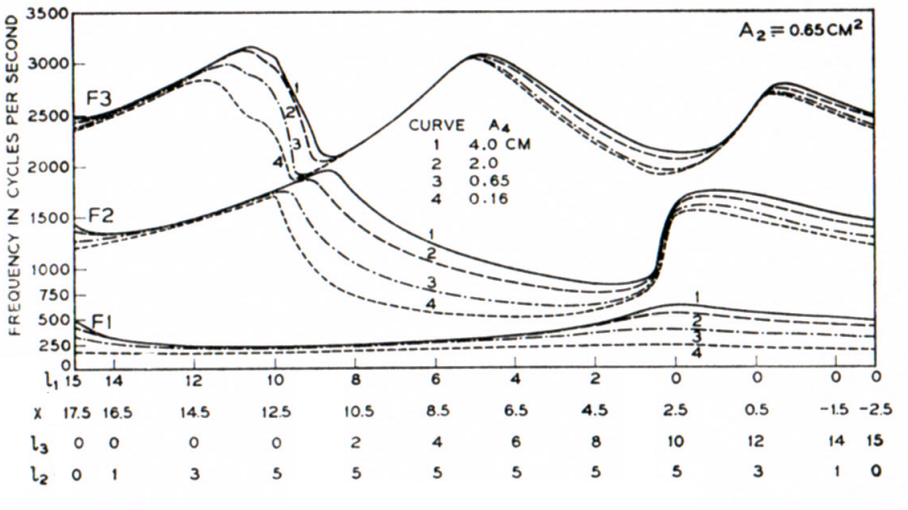
3.1.3 Acoustics of the vocal tract

The vocal tract is an acoustical tube with nonuniform cross-sectional area. It functions as a variable acoustic filter, shaping the spectrum of the glottal excitation produced by the vocal cord vibration. The spectral shaping is achieved by local constrictions along the tract, resulting a cascaded sections with different lengths and cross sectional areas. Different combinations of sections provides a unique resonant system that filters excitation signal to a phonemes as a part of a understandable speech.

The vocal tract starts above the vocal cord constriction at the top of the trachea and it is terminated at another end by the lips and nostrils. This part of speech production mechanism is marked in figure 3.1a. Above the vocal cords there are two cavities that are separated from each other by an epiglottis, that is plate shaped cartilage laying behind the tongue. Epiglottis acts as a valve, obstructing the path to trachea when swallowed. Epiglottis doesn't take part in articulation process. Another gate in vocal tract is a velum, that controls the opening between oral cavity and nasal cavity during the articulation. Most of the job in vocal tract shaping is done by a tongue. It is a large system of muscles connected in front to the lower jaw and in back to bones of the throat and head. The vocal tract is terminated by the lips, that forms the last constriction point before the sound radiation takes place. The length of the vocal tract of an adult male is about 17 cm and its cross sectional area can vary from zero to 20 cm^2. (15)



(a) Illustration of four-tube model.



(b) The variation of first three formants for a range of parameters for a constant tongue hump constriction area.

Figure 3.4: The effect of different parameters on three first formants (F_1,F_2,F_3) on four-tube model of the vocal tract. The different curves are for four different mouth areas ($A_4 = [4.0, 2.0, 0.65, 0.16]$ cm^2). Constant quantities are $A_1 = A_3 = 8$ cm^2, $l_1 = 1$ cm and $A_2 = 0.65$ cm^2. Figures adapted from (15, p.81)

One approximation of the filter characteristics of the vocal tract given by Flanagan is "four-tube, three-parameter approximation of vowel production" (15, p.80-82). The approximation relies on dividing the vocal tract to four parts, that are approximated as tubes of different cross section areas and lengths (figure 3.4a). The four parts are back pharynx cavity (A_1,l_1), the tongue hump constriction (A_2,l_2), the forward mouth cavity (A_3,l_3) and the lip constriction (A_4,l_4). In figure 3.4b the effect of the different parameters to three lowest formants frequencies are shown. A summary of four-tube analysis according to Flanagan (15, p.82) implies that vowel articulation can be described by three parameters. These parameters are the distance from the glottis to tongue hump constriction, x; the size of the tongue constriction, A_2 and a measure of lip-rounding (area to length ration of the lip constriction, A_4/l_4). It must be noted that this analysis is not generally sufficient to describe the consonants or nasal configurations.

3.1.4 Speech radiation

As most of the speech is radiated from the mouth, the main focus is of this chapter is kept on the mouth radiation. The oral tract is terminated at the lips, that are the outermost and visible part of the mouth. By changing the orientation, shape and opening of the lips, a last constriction point of the vocal tract influences in two important factors in speech production; 1) a vocal tract termination impedance and thus the filtering characteristics of the vocal tract (sec.3.1.3), 2) an orifice from which the sound is radiated to a surrounding air. The directivity characteristics of head is highly frequency dependent. The overall effects on sound spectrum radiated from the human mouth varies as a function of angle and distance. The attenuation in function of the distance is due to the attenuation of the transmission medium, and the rate of 6 dB per doubling the distance holds in air for distances more than 10 cm from the mouth (16). When the sound energy radiation takes place at the mouth, the geometry of the mouth, head and torso relates how sound is spread spatially around the speaker. The same approximations of the geometry that were made for the mouth impedances, doesn't necessarily apply for directivity pattern of human head. For example the simple piston in an infinite plane approximation that quite well suited for a mouth impedance approximation doesn't give result out of $\pm$ 90° range around the mouth axis.

In the frequency range where the wavelength is large compared to the transverse dimension of the vocal tract and mouth opening, an assumption of an approximately uniform and cophasic velocity distribution at the mouth can be made. The mouth can therefore considered as a piston or spherical cap shaped radiator, where all parts of the surface are vibrating in phase. The radiator is baffled in a human head, and despite of all the features the shapes of human head, the common simplification of the geometry used is a use of a spherical baffle with radius of 9 cm for a man (15, p.36). It has been argued in the literature (31) that use of a prolate spheroid geometry in a infinite baffle instead of spherical cell to simulate the human head gives more precise directivity measures. At low frequencies where the dimensions of the head are small compared to the wavelength, the influence of the head to the field is minor. Measurements made by Flanagan in the 60's with a real sized mannequin showed that a simple point source approximation gives good result at low frequencies to all angles, and $\pm$3 dB pressure magnitude agreement up to 4000 Hz in $\pm$60° region (16). The directivity measurement made with reciprocity method (20) agreed well on the overall macroscopic characteristic, such ass low-pass filtering as the angle is increased and lobing effect at the symmetry point in the back of the head, but showed that the baseline for comparison, an analytical solution from simple spherical model differs in more detailed analysis. The effect of the mouth cross-sectional area to directivity features was shown to be the most important parameter especially in near field (20).

The nose radiation takes place with the nasal consonants, when a complete closure is made in front of the oral cavity by lips or tongue, and velum is opened to nasal cavity to provide a main sound transmission path. In case of nasals, most of the sound is

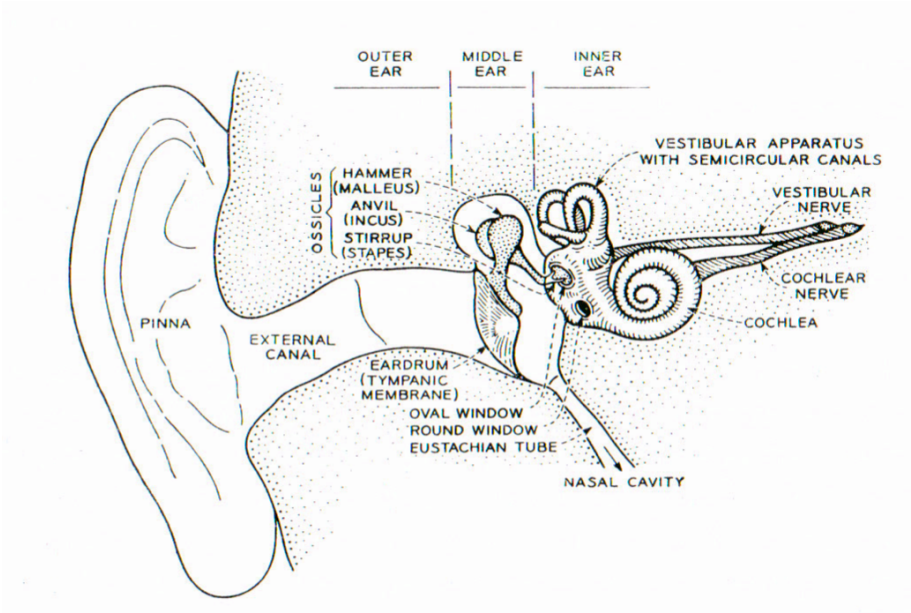


Figure 3.5: Illustration of the human ear showing outer, middle and inner ear. The inner and middle ear structures are shown enlarged for illustrative purposes, so the drawing is not to scale. Figure adapted from (15)

.

radiated from the nostrils, but a closed oral cavity may couple to a radiation as it functions as a side branch resonator in parallel to a nasal cavity. In some languages also nasal vowel are common.

3.2 The ear and hearing

The hearing is a process to receive sound and convert it to nerve impulses by the means of the ear. The nerve impulses are post processed by brains to a perception, in which the heard sound are interpreted and given a meaning. The ears are acousto-neural transducers that are divided into three parts: the outer ear, the middle ear and the inner ear (figure 3.5). The outer ear collects the sound and directs it to a eardrum, after which the middle ear takes care of the transmission of the vibration from the eardrum and fluid filled cochlea in the inner ear. In the inner ear where mechanical-to-neural transduction takes place.

The outer ear consist of the pinna, the external ear canal and the eardrum. The pinna is the visible part of the ears on both side of human head, that surrounds and protects the entrance of the external ear canal. The acoustical function of the pinna is to collect sound and direct it to the ear canal. The pinna affects the incoming sound by applying a frequency and angle of incident dependent filtering and amplification, that are used as a directional cues on sound localization. The pinna affects mostly on localization of sound elevation (19) and front-to-back impression (7).

The external ear canal, or meatus, is a slightly curved transmission canal of about 2.7 cm in length with 0.3 to 0.5 cm^2 cross sectional area in oval-to-circular shape. The meatus is entirely lined with skin and it is terminated by the eardrum (tympanic membrane), that is a thin and stiff membrane with shape of inwardly-directed cone. The ear canal has a resonance behavior of quarter wave resonator, like a uniform tube with one open and closed end. First resonance of the ear canal is around 3000 Hz. The combination of resonances in the cavities of pinna and external ear canal produces a broad peak between 2 and 5 kHz in the amplification introduced by outer ear. Because of the quite uniform diameter around 7-8 mm of the ear canal, a one dimensional plane wave propagation exist up to 25 kHz between the entrance of the ear canal and eardrum. The outer ear is passive, linear and time invariant system.

On the inside of the eardrum is an air-filled middle ear cavity which contains the ossicular bones as shown in figure 3.5. The ossicular bones are three tiny bones that are called malleus (hammer), incus (anvil) and stapes (stirrup). The function of these bones is to present an impedance matching between the air at the outer ear and liquid in the inner ear by conversion of the external sound pressure to a fluid volume displacement. The malleus is attached to a eardrum, and stirrup is attached to the oval window of the inner ear, and the incus connects the two. The overall pressure transformation ratio from eardrum to oval window is slightly frequency dependent but has a value around 12 - 18 to one (33, p.100). The middle ear structure provides also a protection against loud sound. The protection is served by the control of the muscles that can damp the vibration of the eardrum. The middle ear is also connected to the outside world by the way of the eustachian tube, which permits the pressure equalization between the middle ear and atmosphere.

The inner ear is composed of the vestibular apparatus, the cochlea and the auditory nerve terminations. The vestibular apparatus and associated organs are used for balance and sensing orientation and are not associated with hearing procedure. The mechanical vibrations are converted to a neural impulses in the cochlea. The cochlea is shaped like a snail shell and it is communicating with the middle ear via the round and oval windows. The simplified illustration of the straightened cochlea is shown in figure 3.6 and a cross sectional view of it in figure 3.7. The cochlea is divided from the middle by a basilar membrane and Reissner's membrane into a two tubes; the scala vestibuli and scala tympani. Those passages are connected to each other by the opening called helicoterma at the far end of cochlea. The vibration produced by stapes at the oval window travels in the fluid filled cochlea along the scala vestibuli to the other end, where the wave passes trough helicoterma to scala tympani. As the wave travels along the scala tympani, the vibration couples to a basilar membrane. The organ of Corti and tectorial membrane runs next to basilar membrane. The organ of Corti is attached to basilar membrane, and cab between the organ of Corti and tectorial membrane is filled with hair cells. Hair cells transform the vibration of the basilar membrane into electrical signals that are transmitted via auditory nerves to brains.

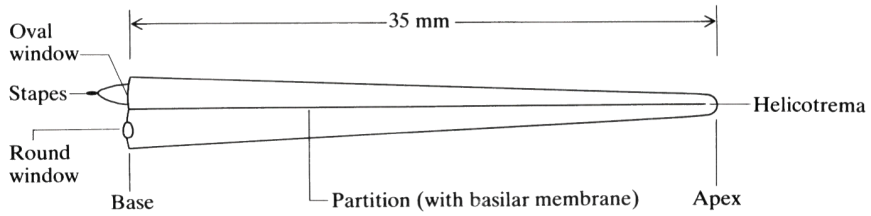


Figure 3.6: Simplified diagram of the cochlea as it would appear if unwound. Figure adapted from (25)

.

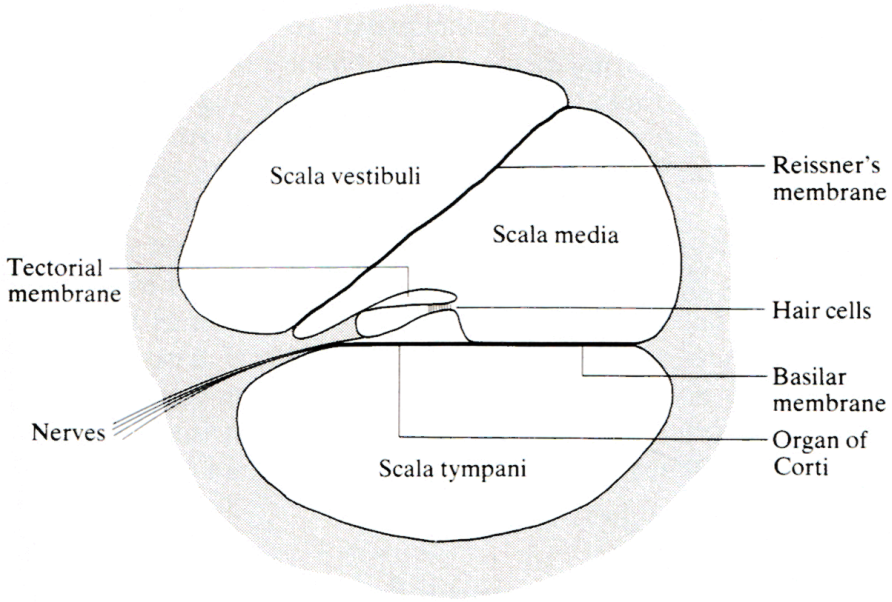


Figure 3.7: Cross section of one turn of the cochlea. Figure adapted from (25)

.

The basilar membrane is stiffer and lighter on the other end and more compliant and massive at the other end. The different parts of the basilar membrane together with its surrounding liquid has a different resonant behavior due to different mass and compliance, thus the basilar membrane vibration has frequency selective behavior.

Auditory masking is an phenomena of sound perception disruption due to a presence of another sound. Auditory masking can take place in frequency or time domain. A loud noise can mask another sound, and thus in a noisy environment a speech transmission and perception can become disturbed (35). The detailed description of complex phenomena of masking is out of scope of this thesis.

Chapter 4

Methods of flow generation

The measurement of the effects of the wind on sound needs ideally an adjustable, predictable and constant flow conditions in a large enough anechoic space without extraneous noise. These requirements are not easy to fulfill. The real life wind situation is hardly possible to achieve at the laboratory conditions, and at the other hand the actual wind conditions outside are not predictable nor constant for repeatable measurements. The wind-induced noise and other factors on the performance of microphones and windshield has been investigated in controlled environments. There are several methods used for flow generation that are introduced in this section, partly after Brixen (10).

4.1 Wind generators

General feature for wind generators is the use of electrical motor to rotate a bladed wheel to displace air. The displaced air is optionally ducted to an outlet from where the flow is transmitted to a surrounding. By shaping the duct the nature of the flow can be modified by some degree. The initial volume velocity can be altered only by changing the rotational speed or modification of the blades. The flow speed can be increased by reducing the duct area focusing it to a smaller region.

The non-ducted wind generator (figure 4.1) described by Olson in (24) was used at the RCA for measurements of wind response of microphones and effectiveness of wind screens (8). The wind generator consisted of large rotating wheel with blades, that was driven by adjustable electrical motor. The wind generator was able to produce almost any practical value of wind velocity and pulses of air. The wind velocity could be adjusted by the speed of rotation. The behavior of the wind could be changed by altering the amount and angle of the blades. The measurements with the wind generator could be done in anechoic chamber without present of the other noise sources than the wind generator itself. Although the conditions aren't defined precisely regarding the balance between the laminar flow and the turbulence, the

measurements with the wind generator showed good agreement compared to the outside measurements in real wind conditions (8).

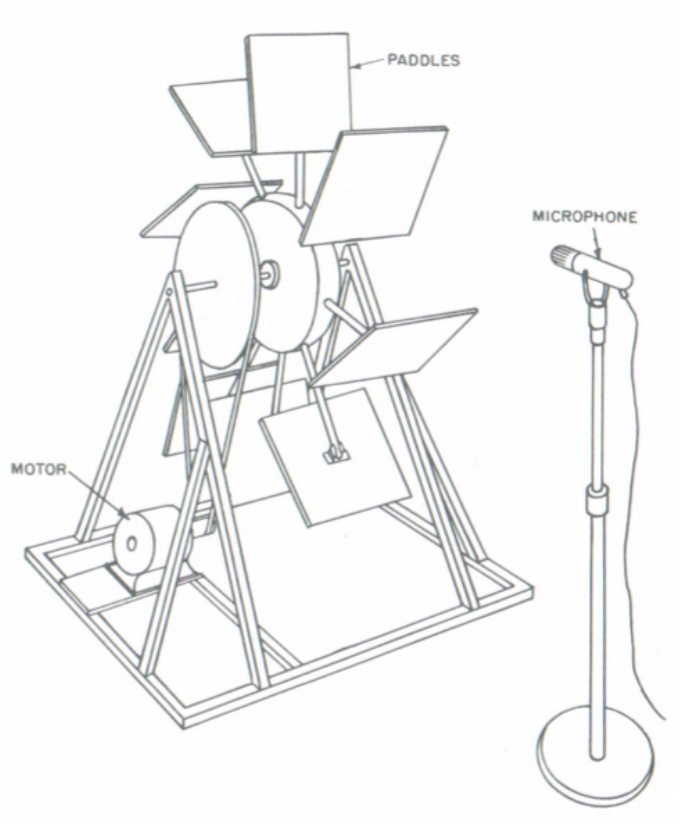


Figure 4.1: A wind generator developed by RCA for obtaining the wind related microphone measurements. Figure adapted from (24)

More recently, two different ducted wind generators are described in the IEC 60268-4 (21) for measuring the wind sensitivity of microphones. Both generators relates on mechanical fans with electrical motor and blades for air displacement. First generator is radial fan with short flow duct and another has a longer duct and axial fan. The shorter with radial fan (figure 4.2) has been used in many institutions for obtaining comparative results. When the results between the described wind generator are compared to other type of wind generators, major differences are shown, so the results are just comparable if the same kind of wind generators are used. This type of wind generators produces a noise that cannot be neglected. (8)

A silent flow generating system has been constructed in NAL[1] for measuring the acoustical properties and aerodynamical performance of acoustic louvres. The overhead diagram of the system is shown in figure 4.3. The system relies on effective silencers and stabilizing the turbulent airflow by using an air plenum in front of the exhaust duct, that is terminated in the reverberation room. The noise level generated by the fan and turbulence in the production of airflow of 2 to 8.5 m^3/s was

[1]National Acoustic Laboratories, Australian Hearing

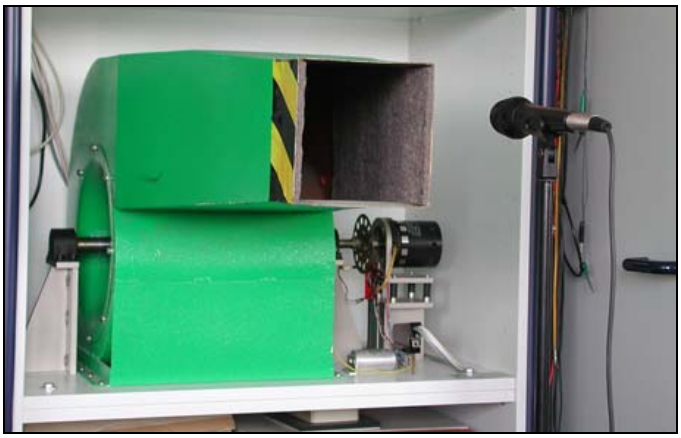


Figure 4.2: Wind generator with radial fan, as per IEC 60268-4. Figure adapted from (29)

between 40 to 60 dB(A) at the entrance to the reverberation room testing chamber. More detailed description of the construction and performance of the system is given in (14). The described system was used in the study of wind noise within hearing aids conducted by Zakis (34).

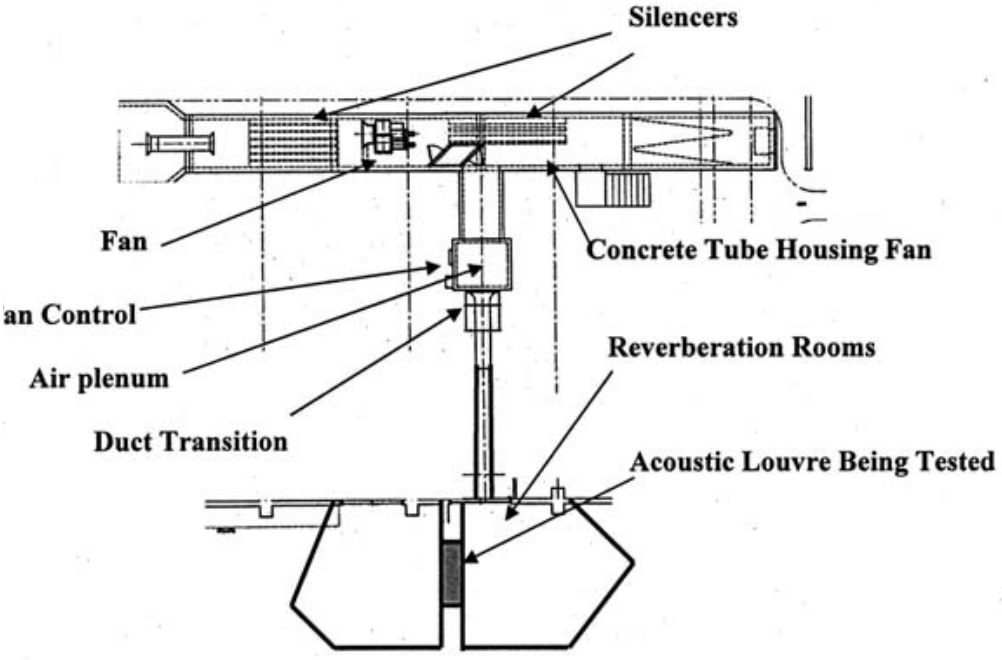


Figure 4.3: NAL testing facility layout. Figure adapted from (29)

4.2 Wind tunnels

The wind tunnels have been the major testing and verification tool in aerodynamic research from the end of 19th century. The wind tunnels have developed from a simple box with air blown trough it by a steam-powered fan to a supersonic tunnels with various forms and scales in between them. The wind tunnels differs from wind generators so that the device under test is placed inside the tunnel, rather than outside of the duct end termination in the surrounding air. Two common designs are open circuit wind tunnel and closed circuit wind tunnel. Open circuit wind tunnels are ducts with inlet and outlet, and the air is flowing trough the tunnel. As shown in figure 4.4, open wind tunnels are usually straight ducts with the fan located at the exhaust end, and air is flowing in from the other end trough straightening vanes for smooth and uniform flow. Closed circuit wind tunnels are return-type configurations where the same air is rotated in the system, as shown in 4.5. The wind tunnel design aims to provide the optimal and controlled flow conditions in test section. (30)

A wind tunnels practical for sound related research needs to have sufficient cross sectional area of test section, low noise and ideally highly absorbing walls. If only wind noise is studied, the dimension requirements can be relaxed by some degree and only controlled low noise airflow is required. If the effect of flow on sound radiation and transmission is studied, the presence of the reflecting walls becomes a problem. It has been noted by Brixen (10) than even the best tunnels do become noisy at very high wind speeds.

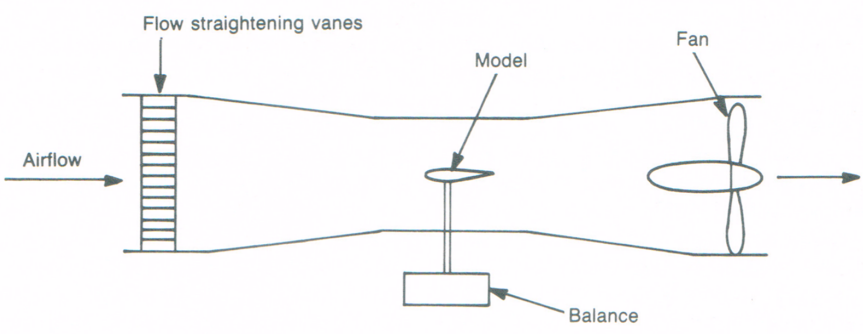


Figure 4.4: Diagram of an open circuit wind tunnel. Figure adapted from (30)

4.3 Other methods

A rotating boom can be used to move the device under test (DUT) in anechoic conditions with a constant velocity. The produced flow is more or less perfectly laminar. If turbulent flow is needed, some object can be placed in front of DUT. The rotating boom has been used to study the wind noise in microphones and windscreen analysis by Brock (11) and Hill (18). Rotation of the boom can create large scale vortices inside a room that may effect the results on following rounds. Also the

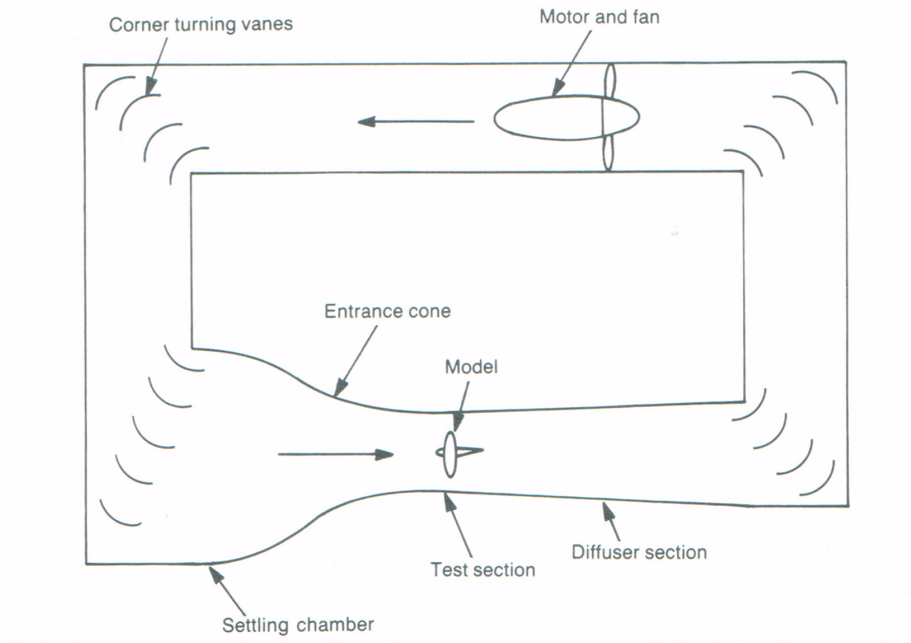


Figure 4.5: Diagram of an closed circuit wind tunnel. Figure adapted from (30)

effects near the room boundaries should be taken into account. The system can be made quite silent if the motor is installed in different room and the dampening of mechanically vibration is taken care of.

As summarized by Brixen (10), windshield performance testing has been carried out by "pole out of car window" method. The microphone together with windshield is mounted on a pole, and the whole system is fixed out of the car window and the car is introduced in a motion. The method is vulnerable for inaccurate speed control, turbulence around the car mixed with possible real wind conditions and noise from tires and engine. The method is used for example in measurements of windscreen performance for sound velocity sensors (23).

Realistic wind conditions can be achieved with outside measurement. The outside measurement offers valuable information of the device behavior under real wind conditions. The natural wind cannot be controlled, but the nature of the wind can be monitored by using anemometer to record the wind velocity and direction.

Chapter 5

Measurement methods

In chapter 4 the flow generation methods was discussed. In this section the flow generation method and the rest of the measurement system used in this research is described.

5.1 Pendulum system for flow generation

A method to produce a flow around object in anechoic chamber with as low background noise as possible was designed and built. The basic idea was to create a pendulum whereby a sound sources or receivers could be introduced into a motion with velocity higher than 10 m/s . Such system is free of noise that would exist in case that a flow would be produced by means of mechanical fan. In fans the noise is produced by the power source of rotation, such as an electrical motor, and the noise produced by flow around the blades and frame. Another important factor is that by using a pendulum system in an anechoic chamber, the absence of reflecting surfaces is maintained with high degree.

The pendulum system was installed to hang from the ceiling of the anechoic chamber (figure 5.1). The distance between the wedged on the ceiling and the walking support wire grid was limiting the dimensions of the structure. The supporting top frame, the pivot swing-axle, the first part of the main pendulum frame and a movable horizontal bar was made out of rigid steel. The pivot swing-axle was attached to a top supporting frame with two ball bearing units as shown in rendered drawing in figure 5.2. The movable horizontal bar acts as a support for the additional mass to shift the moment of inertia to optimal position in a frame. The rest of the light body of the pendulum was made of an aluminum tube with 25 mm outer diameter and 2 mm wall thickness. The tubes was attached to each other and to the top part of the frame with M8 bolts and nuts. The light part of the pendulum frame was designed to be as rigid as possible. Other important criteria was to minimize the mass and the area of reflecting surfaces of the frame and to get the initial angle of

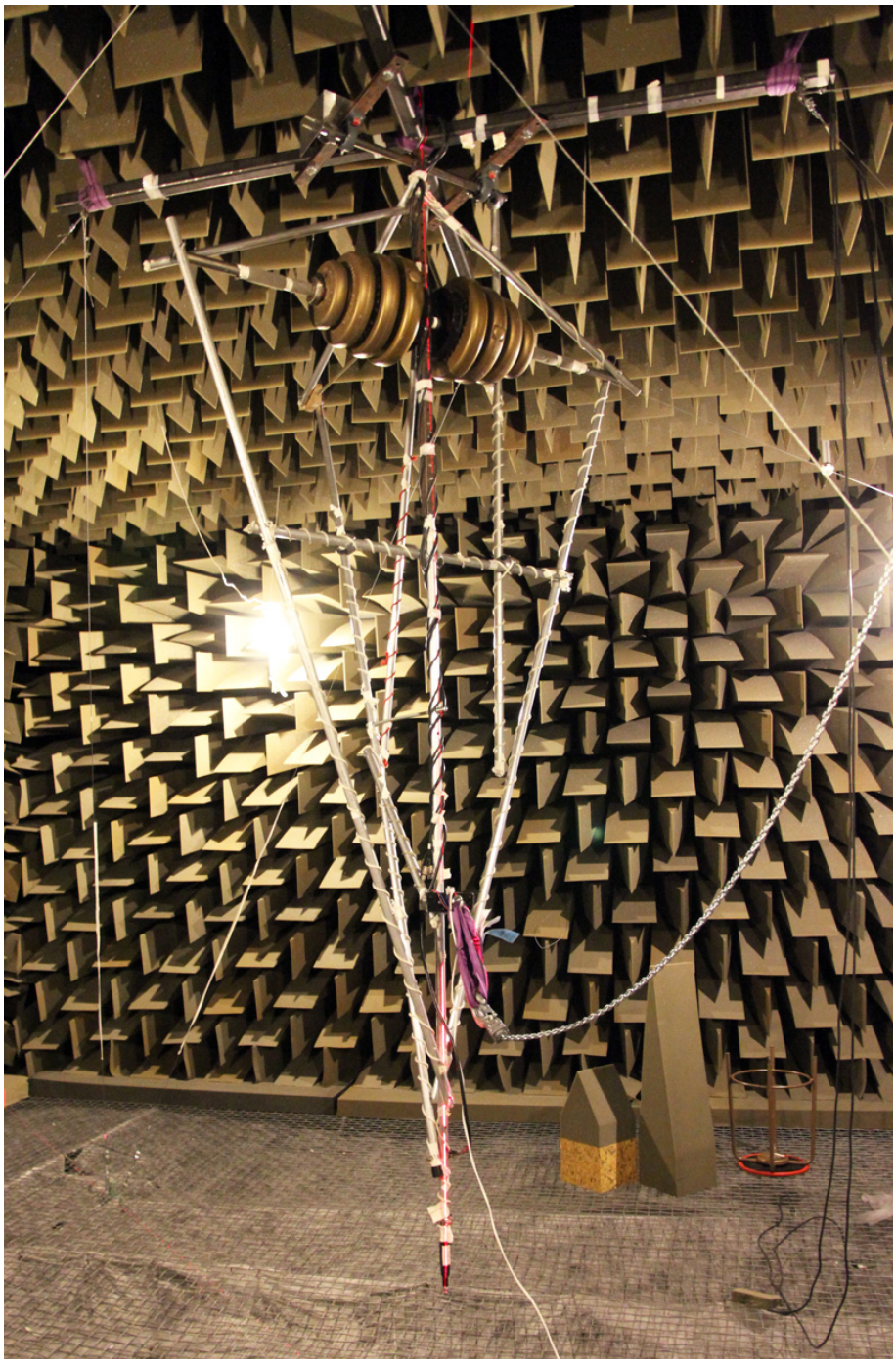


Figure 5.1: The pendulum system under an alignment and guy-wire adjustment without a measurement source or receiver attached. A loosely hanging chain on the right is the lifting chain and a white cable at center bottom is for the magnet release current.

release as high as possible. Attention was also paid into an aerodynamical properties of the frame to minimize the drag and turbulence induced noise.

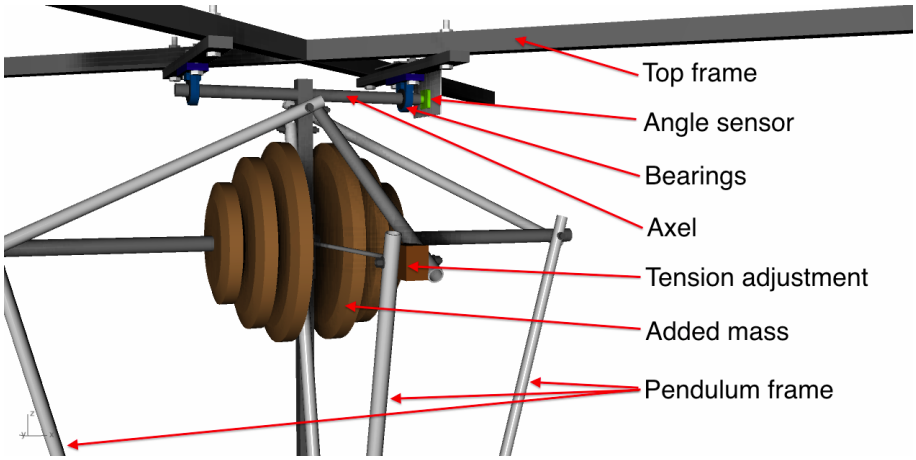


Figure 5.2: The rendering shows the design of the top frame and attachment of the pendulum to the frame together with some other details. The tension adjustment together with another similar one lower was used to prevent the ending of the frame in acceleration. The actual implementation is similar to this rendering.

The pendulum system was lifted up together with the top support frame by means of electrical chain hoists, and then hung up with wires trough holes of the ceiling of anechoic chamber. The pendulum system was guy-wired from its top support frame down to a fixing points to hold the top support frame steadily in balance. The initial angular displacement of the pendulum was achieved by the means of electrical chain hoist and strong electrical magnets. The magnets could easily hold the weight of the pendulum and the magnet field could be removed with 2.2 A current for precise release of the pendulum into a swing. The current was supplied from the power supply units through a purpose built high power switch. The switch could be controlled with 0/5 V control signal that could be supplied by means of mechanical switch, programmatically controlled micro-controller or another hardware. The high power switch was implemented with a darlington transistor pair for high current gain. A flyback diode was used to prevent the sudden voltage spike during the switch off across the inductive load of a coil. The angular displacement θ was monitored with non contact hall effect rotary position sensor. Sensors position marker magnet was attached concentrically on the end of the pivot swing-axle and the sensor housing was attached on the top support frame. The sensor outputs an analog voltage level related to the supply voltage and the angle between the sensor and the rotating position marker. The output was connected to a measurement system to record the level. The output level in function of the displacement angle was calibrated prior to measurement sessions. The angular velocity ω could be determined by time derivation of the angular displacement $\omega = \mathrm{d}\theta/\mathrm{dt}$.

5.1.1 Velocity of pendulum

The pendulum is called a physical pendulum, if it has an extended mass distribution along the body rather than all the mass would be concentrated at a single point, what is the case of simple pendulums. The pendulum equation for physical pendulum is

$$I_p \frac{d^2\theta}{dt^2} + k\frac{d\theta}{dt} + mgl\sin\theta = H \tag{5.1}$$

where I_p is the moment of inertia about the pivot point, k is the damping coefficient, l is the distance from the pivot to the center of mass, m is the total mass of rotating structure, g is the acceleration due to gravity, θ is angular displacement from the vertical rest position and H is the applied external torque (26). The objective was to find a optimal location and mass for added mass to maximize the angular velocity of the pendulum. As a precise solution of the motion of the pendulum is out of scope of this thesis, an approximations and linearization of equation 5.1 can be made. Usual linearization for the equation 5.1 counts on the approximation that $\sin\theta \simeq \theta$ if θ is small enough (26; 4; 5). With such linearization the differential equation has straightforward solution and the motion of the pendulum can be described as a simple harmonic oscillation, and the angular velocity could be determined. But with large amplitude displacement the approximation ($\sin\theta \simeq \theta$) does not hold, and the natural period is amplitude dependent (5). The maximum angular velocity as a function of the moment of inertia and the center of the mass can be optimized by the use of energy conservation theorem. As an approximation the damping effect of the friction due to a bearings at the pivot and the aerodynamical forces was diminished by setting $k = 0$. Applying the first integral of motion to the equation 5.1, and let $H = 0$ as it is the case, we get

$$\frac{1}{2} I_p \left(\frac{d\theta}{dt}\right)^2 + mgl(1 - \cos\theta) = \text{const.} \tag{5.2}$$

The first term of equation 5.2 is kinetic energy and second term is potential energy (4). In conservative system the sum of kinetic and potential energy does not change over time and is constant. The constant term is the potential energy of the pendulum in rest (4) . The equation can be written in a following form to solve the the angular velocity ω

$$\omega = \frac{d\theta}{dt} = \sqrt{\frac{2mgl(1 - \cos\theta)}{I_p}} \tag{5.3}$$

An additional mass of 69 kg was chosen as a reasonable compromise between the gain in angular velocity and the stress it introduces to the structure. The location of the additional mass was chosen to provide maximum angular velocity when 2 kg mass

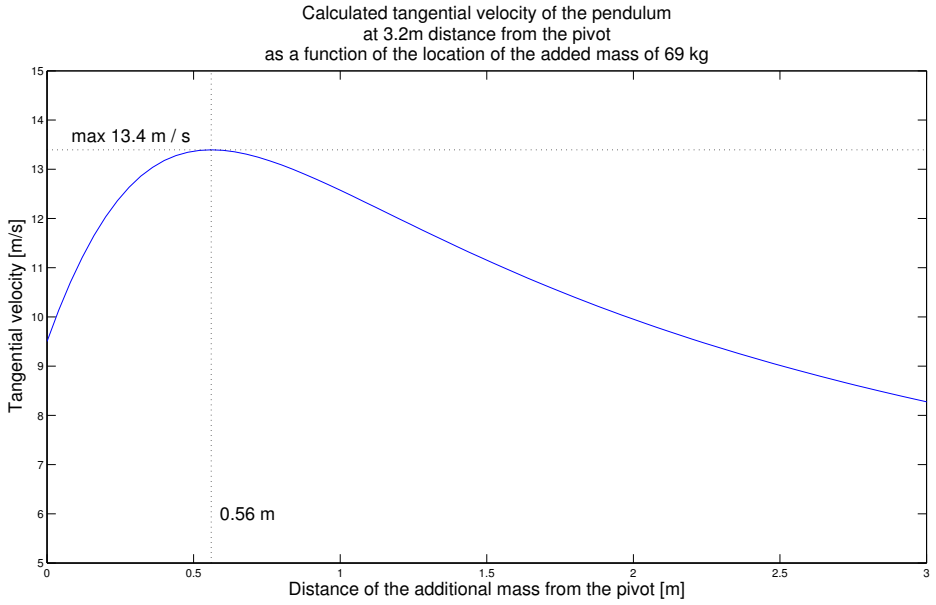


Figure 5.3: Tangential velocity of the 2 kg point mass at the end of the pendulum (R=3.2 m) calculated as a function of the location of the additional mass of 69 kg. The mass distribution of the frame of the pendulum structure is taken account.

distribution from the device under test was located at 3.2 m from the pivot point. The mass distribution of the pendulum frame was taken account. The calculation of the angular velocity according to equation 5.3 as a function of the location of the additional mass is shown in figure 5.3.

5.2 Data acquisition system

The data acquisition and measurement control system was put in practice over a National Instruments (NI) hardware controlled by LawView software. Maximum of six audio microphones could be recorded simultaneously in sync with available NI hardware. In directivity measurements, where more than six microphones was needed, a parallel system consisting of a RME Fireface 400 audio interface and Matlab program was used. The synchronization of the two measurement systems was done by the means of a high frequency synchronization signals, and data from the two system was time aligned in post processing phase. All the post processing and data-analysis was done in Matlab environment. The measurement data was transferred from the LabView environment to Matlab by means of comma separated ASCII-files.

All signal output and input events were started by a trigger that activated when the pendulum exceeded a set angular displacement threshold. The voltage level corresponding to angular displacement was known from the calibration procedure.

As the hardware didn't support analog voltage triggering, a comparator circuit was implemented to generate a digital trigger signal according to set voltage threshold. The hardware triggering was needed, because if the decision of the exceeding of the certain voltage level was done in a software level, the delays due to a software, operating system and hardware interaction didn't remain stable and precise triggering could not be implement. The pendulum was released from its initial position by giving +5VDC control pulse to the magnet release circuit. When parallel Matlab based acquisition was needed, a Matlab controllable micro-controller was used to output the control pulse for synchronization purposes. Otherwise a separate mechanical switch was used for controlling the magnet release circuit.

The common problem with audio interfaces are the disability to record DC-voltage levels or very low frequencies, that was needed for monitoring the angular displacement. In the first phase of measurements all the data was recorded with audio interface and Matlab. The angle sensor output was used as control voltage for purpose designed and built voltage controlled gain circuit, by which a high frequency sine signal was amplitude modulated in function of the angular displacement. With such procedure angular displacement could be monitored in synchronous with microphones by extracting the modulation envelope from the carrier wave in post processing phase. The need for precise angular displacement based triggering of the signal output and input events evoked the re-arrangement of measurement platform to better suitable hardware.

5.3 Simplified artificial speech source

For measuring the effect of the flow on radiation from the mouth, a simplified artificial speech source was built to be used together with pendulum. For the measurement purposes, the geometry of a human head and vocal tract was simplified. The design is shown in figure 5.4. A hard plastic spherical cell with 15 cm radius was chosen as a baffle for the artificial speech source. To model a neutral vocal tract, a straight tube with hard walls was flush mounted inside a sphere. The length of the tube was 17.5 cm and the diameter was 2.6 cm. A moving coil cone type loudspeaker driver with a 70 mm nominal diameter was coupled to the tube trough a small compression chamber. The back of the loudspeaker unit was enclosed in a damped, rigid, airtight enclosure. The tube mounting flanges was designed and made for easy change of tubes and adding sensors into a throat of the tube. The spherical cell was filled with a foam to reduce the ringing of the cell introduced via the mechanical and acoustical coupling from the loudspeaker driver. The magnitude response of the simplified artificial speech source is shown in figure 5.5. The weight of the artificial speech source was tried to maintain as low as possible, so that the forces acting on a sphere when accelerated and the effect on center of the mass of the pendulum construction would be minimal. The loudspeaker driver had a neodymium magnet assembly. The weight of the artificial speech source with all the fittings was 2.21 kg.

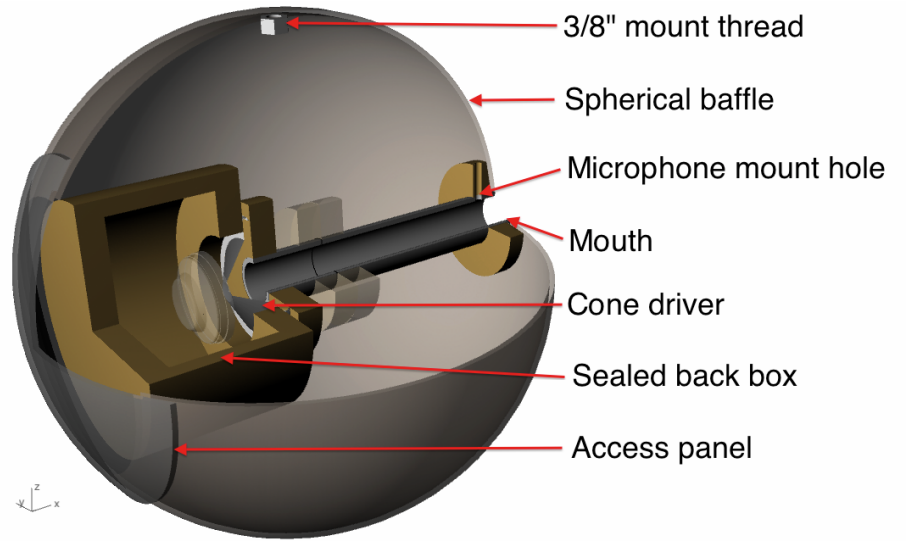


Figure 5.4: The rendering shows the design of the artificial speech source. The frame of the cone loudspeaker is not shown as a whole for better visibility. The actual implementation is similar to this rendering.

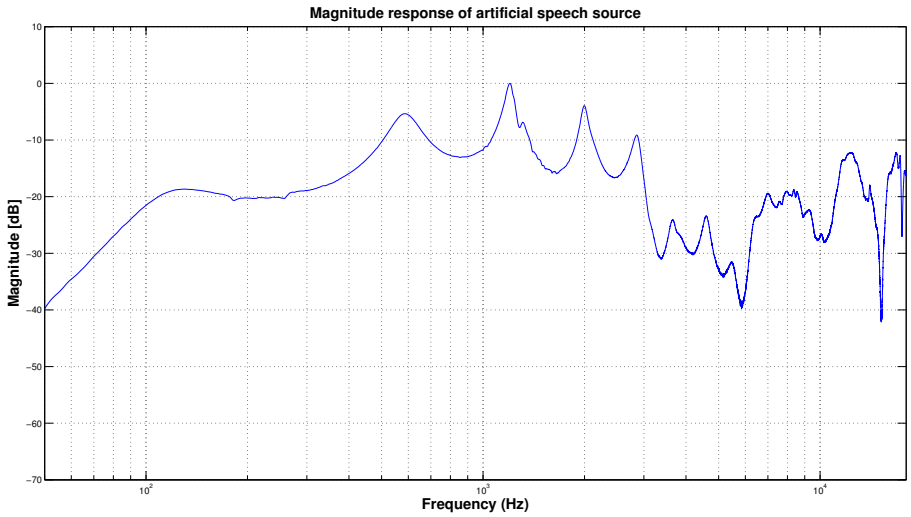
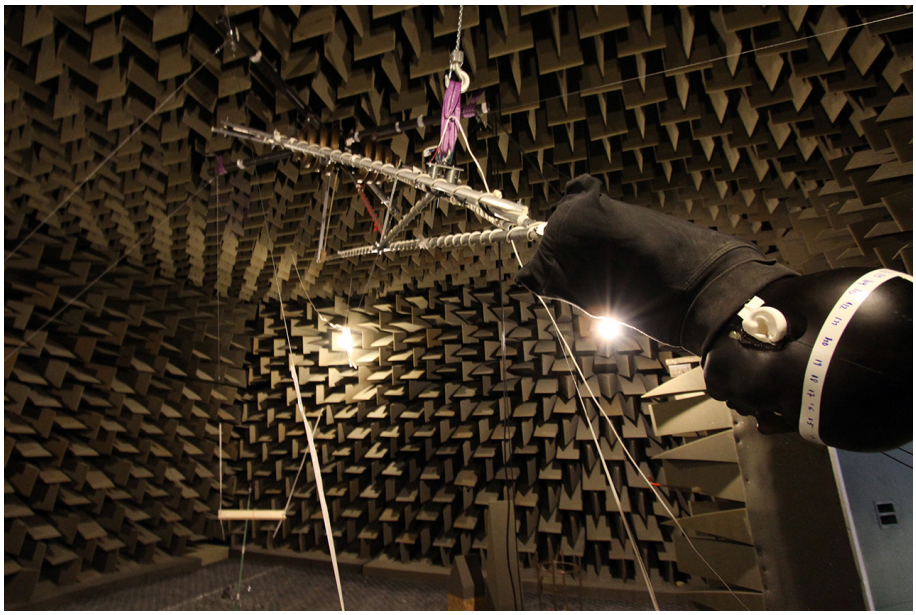


Figure 5.5: Magnitude response of the simplified artificial speech source measured 10 cm in front of the mouth aperture.

5.4 Dummy head for flow induced binaural noise measurements

For recording a flow induced binaural noise, a dummy head with adjustable ear canals (DADEC) (17) was used for blocked ear canal recordings. The DADEC utilizes a real size head and torso together with acoustically realistic artificial pinnae and ear canal. A rigid mounting point for pendulum usage was added inside DADEC for stable mounting and easy angle adjustment. A Knowles FG 23329 miniature



(a) Dadec raising up with the pendulum.

(b) In-ear miniature microphone inside the ear canal.

Figure 5.6: Dummy head with adjustable ear canals (DADEC) for blocked ear canal measurements in pendulum system.

pressure microphones mounted in the center of soft ear plugs was used to measure the pressure at the entrance of the ear canal as shown in figure 5.6.

Chapter 6

Measurements and results

6.1 Measurements

Several different measurements was done by using the measurement setup described in chapter 5. The aim of the measurements was to find out whether flow have measurable effects on the radiation and transmission of sound from the artificial speech source. In contrast to the behavior of the artificial speech source, binaural noise was recorded with dummy head in similar flow conditions. The set of measurements should give a good measure of the acoustics of speech communication in present of the flow.

The measurement principle in all measurements is to record a sound pressure in a specific location and analyze the recorded data. In case of a sound production and transmission, a known signal with specified characteristics was played back with the artificial speech source. Two recordings were made; a reference was recorded without the flow and then another recording with the flow. Only the presence of the flow differs between the recordings while other parameters are kept constant. The recordings were analyzed and compared in frequency domain.

By measuring the sound pressure response both inside the vocal tract and on different angles further away from the mouth aperture, the effects of different parts of the transmission path was discretized and their importance on the total effect could be evaluated. After a short description of the used measurement signals, different measurements are explained in this chapter with some discussion about the effects that were looked after.

6.1.1 Measurement signals

Several different continuous signals were tested and used as excitation signals in the measurements with artificial speech source. The most natural choice for excitation signal would utilize same characteristics than glottal excitation for voiced speech production. Such signal was created by looping one cycle from the glottal source

signal for Finnish vowel /a/ used by Takanen & all (32). In their work the glottal source signal was separated from the anechoic recording of an utterance of the vowel by using glottal inverse filtering algorithm, Iterative Adaptive Inverse Filtering (IAIF) (1; 2). The original waveform is presented in figure 6.1 and the properties of the generated glottal excitation measurement signal is shown in figure 6.2.

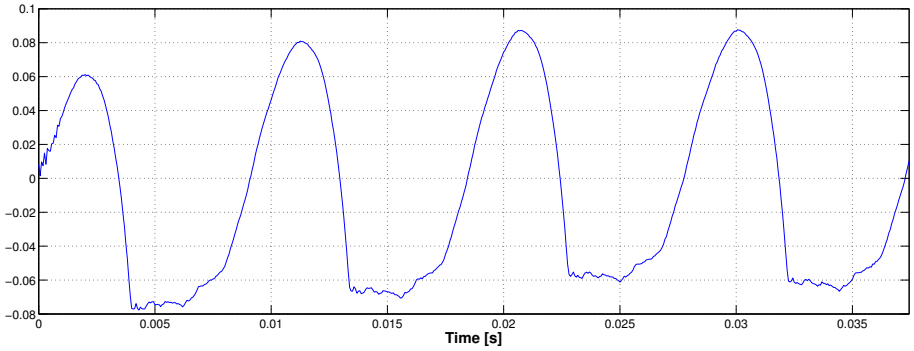
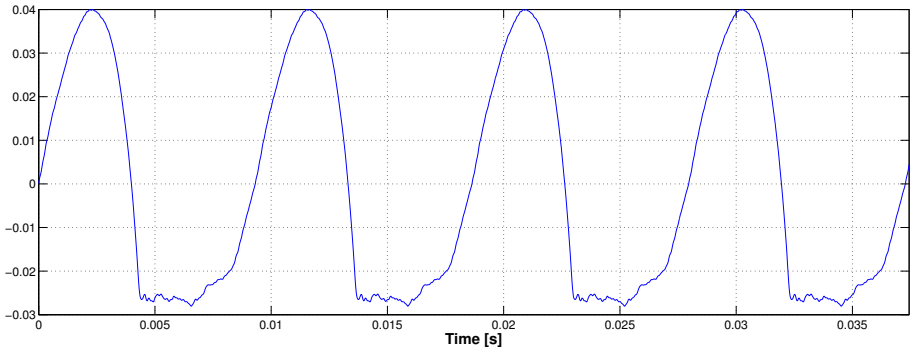
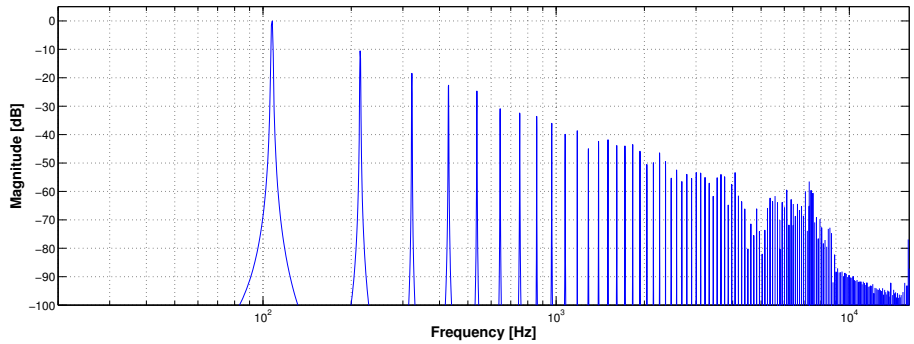


Figure 6.1: Glottal source signal for Finnish vowel /a/ after Takanen & all



(a) Waveform



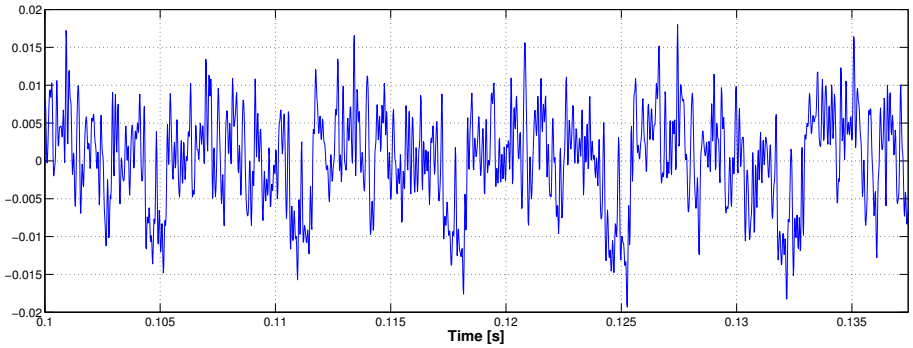
(b) Magnitude response

Figure 6.2: Generated glottal excitation measurement signal.

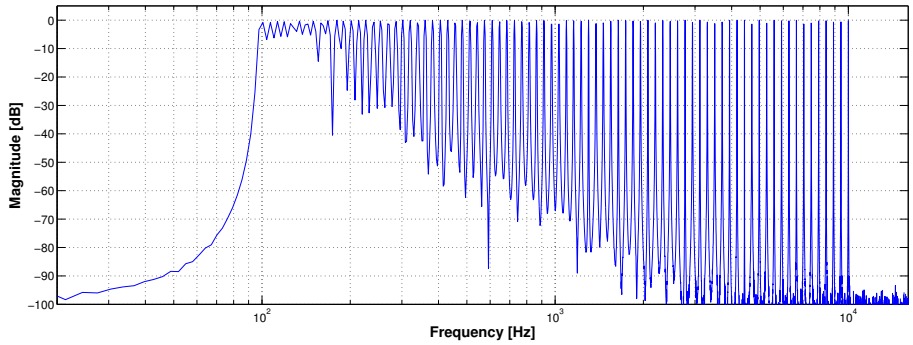
Due the low pass characteristics of the glottal source signal, excursion problems at the low frequencies and poor signal to noise ratio at the higher frequencies limited the usability of this excitation signal. To overcome a limitations of the generated glottal signal, a multi-tone signal was used as another excitation signal. The multi-tone

signal was made up of 80 sine components. Each component had an equal amplitude and components were logarithmically distributed between 100 Hz and 10 kHz as shown in figure 6.3. The multi-tone signal had a flat spectrum over a frequency range of interest. Compared to the generated glottal source excitation, the signal component distribution around the first resonances were denser, that made it easier to identify slight changes in vocal tract filter characteristics of the artificial speech source.

The differences of the harmonic glottal excitation measurement signal and non-harmonic multi-tone signal was studied in preliminary measurements without finding any effects. Next introduced measurement of the effect of low-frequency tone on components around formant frequencies was used to confirm the usability of non-harmonic measurement signal in further measurements.



(a) Waveform



(b) Magnitude response

Figure 6.3: Multitone signal used as an excitation signal in the measurements.

6.1.2 Effect of low-frequency tone to radiation of high frequency components from mouth in presence of flow

In this set of measurement it was studied whether the amplitude of a low-frequency tone has an effect on the transmission of the single or multiple tones located at the resonance or antiresonance frequencies of the artificial speech source. Also the effect of the component on harmonic and non-harmonic frequencies was examined.

Excitation signals were generated by adding the test frequency or frequencies and the low-frequency tone tone. Test frequencies was selected either to match or differ the resonance frequencies. In case of multi-tone test, amplitudes of the tones were equal. The amplitude of the low-frequency tone was changed in each case to determine the existence or significance of the possible effect.

Three different test signals were used. First signal had single tone at 2853 Hz. That frequency was the last resonance frequency before the notch in 3-6 kHz region on the magnitude response of the artificial speech source (figure 5.5). Second test signal was multi-tone signal with four frequency components corresponding four first resonances; 597 Hz, 1214 Hz, 2000 Hz and 2853 Hz. Third test signal was dual tone signal of two antiresonance frequencies 812 Hz and 2450 Hz. With all three test signals, four different amplitudes of the low frequency tone were used; -26 dB,-14 dB, -6 dB, 0 dB relative to tone amplitudes. Fifth case was reference without the low-frequency tone. The frequency of the low-frequency tone was 100 Hz.

Three different angles of attack were studied. Angles were 0°, 90°and 180°. The test signal was played back trough the artificial speech source and recorded with a miniature microphone. The microphone was mounted on the pendulum frame, 1.2 meters away from the speech source towards the axis of rotation. A reference case was first recorded without the presence of flow. Next the pendulum was raised up, and measurement program was started. The pendulum was released into the swing. The program started the synchronized playback and recording when the trigger signal was supplied by the angle threshold trigger circuit. Microphone and angle sensor data was recorded. A spherical foam type windshield was used together with the mic to suppress the effect of the flow on microphone capsule.

Data was analyzed in frequency domain. A sequence of 4096 samples was cut from the recording for a DFT analysis. Amplitudes of different signal components of the test signals was compared to reference measurements.

6.1.3 Effect of flow to pressure inside vocal tract

The purpose of this measurement was to find out whether the magnitude of a pressure would change inside the vocal tract due to flow over the artificial speech source. Measurement point was located inside the vocal tract tube at the distance of 13 mm from the mouth aperture plane. A miniature pressure microphone was flush mounted on the cavity wall. A major effect on pressure response found from this point would indicate a change in radiation conditions at the mouth. As noted in section 3.1.3, the change in radiation impedance would occur as a shift in formant frequencies or change in resonance amplitude or bandwidth.

Both of the described excitation signals were used in this measurement. The multi-tone signal was found give better SNR over the formant frequencies and to indicate more detailed information of the filter characteristics over wide frequency band. In other hand, the glottal excitation signal couldn't show any more results because of

the harmonic nature compared to the multi-tone signal. By the use of multi-tone signal, amplitudes of distinctive frequency components were easy to separate from the background noise.

Examination of the pressure response should be restricted to the frequency range where plane wave propagation takes place inside the vocal tract. The cut-off frequency for the one dimensional propagation in a tube with diameter of a is given by $ka = 1.84$ (13), so a plane wave propagation exist up to 7500 Hz in a 26 mm tube used in the artificial speech source. Higher order radiation modes exist above the cut-off frequency and pressure measured at tube wall is no longer the same across the cross-section plane.

The artificial speech source was rotated in 30 degree steps so that the flow direction related to the mouth aperture was changed accordingly. With each direction, a reference recording without the flow was recorded for comparison.

In post processing phase, a sequence of 4096 samples of the recording was analyzed. The recordings were done with 50 kHz sampling rate. The sequence was cut around the time instant when the pendulum passed the vertical orientation and the tangential acceleration due the gravity was zero. The reference sequence was cut from exactly the same place from the recording of the reference case without the flow. A tapered cosine window with 21 sample wide cosine lobes was used to window the inspection sequence. The windowed sequence was then zero padded and 50k point discrete Fourier transform DFT was applied in order to get a 1 Hz frequency resolution. Magnitude values of the individual signal components of the excitation signal was examined and compared. The results are shown in chapter 6.2.2.

During the analysis sequence of 82 ms, the mean velocity of the artificial speech source was 12 m/s and it travelled 98 cm along the trajectory.

6.1.4 Effect of flow to directional characteristics of artificial speech source

The depence of direction of sound radiation and propagation of the artificial speech source in presence of flow was studied with a polar response measurement. The measurement was carried out with microphone array consisting of seven static microphones placed in 30° interval on an arc of semicircle of 1.9 m radius on horizontal plane. The centre of semicircular microphone array is coincident with the centre of the artificial speech source attached to the pendulum in rest. The line of symmetry of the semicircular microphone array points in to the direction of normal to the plane defined by the trajectory of the pendulum swing. This way the 0° and 180° microphones are on the plane of swing. 0° microphone is on the direction of the flow during the measurement.

The directional data was measured with different angles of attack of 30° multiples. The angle of attack is the angle between the reference line of the body of the source

and the incoming flow, namely the direction of the mouth opening and the forward direction of the pendulum swing. To get directional data over a full circle, two measurements were needed. The second measurement was done with source rotated across the plane of swing defining the wind direction. This way the microphone array was mirrored on the other hemisphere of the source compared to a first measurement case. The mirroring across the plane of swing can be done due to symmetry of the artificial speech source. In case of 0 ° and 180° cases only one measurement is sufficient due to the assumption of even radiation pattern in both hemispheres.

The actual measurements was done by first setting the angle of attack by directing the mouth aperture in the direction of interest. First the reference measurement was done with a pendulum in rest. Then the pendulum was raised in its initial position of about -90 °. The measurement program was started. The program waited the signal from the trigger circuit monitoring the angle sensor data. The pendulum was released in to a swing. When the pendulum had reached the angle of trigger, program started synchronized playback of multi-tone signal (chapter 6.1.1) and recording of the angle sensor data and data from the microphones. After the first measurement, mouth aperture was directed in across the plane of swing for a second measurement of another half space. The reference measurement was done in prior to lift up for flow measurement.

For successful construction of directional response of the moving source from the data captured by static microphone array, a precisely calibrated angle sensors and its exact synchronization to microphone data acquisition was needed. The key factor was to determine the exact location of the source related to the microphones. First, the location and thus time dependent distance from the moving source to every seven microphone locations was determined. Respectively, related direction vectors pointing from source to each microphone, as well as velocity component into those directions was determined as a function of time. The distance data was used to balance the attenuation variation of the recordings in every microphone. The velocity and distance data was used form the time domain mapping function. It relates the time instance of sound wave emission at the source and its reception at the microphone location. By use of the generated time mapping function, the doppler effect could be removed from the recordings (equation 2.4. From the distance and velocity corrected recordings the frequency responses could compared.

The two half space responses are combined in a single polar response in post processing. A high frequency mismatch between the two half space measurements can be seen in polar plots (figures A.19 and A.20). A sequence of 4096 samples was analyzed from the dataset sampled with 50 kHz sampling frequency. The data is presented as a relative gain compared to static case without a flow as a function of frequency in chapter 6.2.3. With this method it is easily shown whether there are gain or attenuation in some frequency regions due the wind in some directions. More detailed set of polar response measurement are shown in A.1.

6.1.5 Flow induced binaural noise at entrance of ear canals

The dummy head described in chapter 5.4 was used to record the flow induced noise levels due to flow. The purpose of the measurement was to compare sound pressure levels between the ears with different flow directions. In order to measure sound pressure levels the microphone sensitivities needed to be checked.

Sensitivity check was done in comparison with G.R.A.S. Type 46AF 1/2" measurement microphone. First the sensitivity of G.R.A.S. microphone was checked by using the 1 Pa at 1 kHz calibrator (B&K Type 4231). Then the G.R.A.S microphone and Knowles FG 23329 miniature microphones were placed next to each other in anechoic chamber. Genelec 8030A loudspeaker was used to produce a 1 kHz from the distance of 1.5 meters. The sensitivity of G.R.A.S microphone was used to determine the sound pressure level at that location, and a sensitivities of Knowles miniature microphones was calculated. The calibration program was set to record the voltage values of incoming signals, so the pressure sensitivity was recorded at Pa/V at 1 kHz value. The noise floor of the microphones was also recorded and it is shown in figure A.21. Noise spectrums were similar between the microphones.

After the pressure sensitivity measurement, miniature microphones were mounted at the entrance of dummy head ear canals, as described in chapter 5.4 . The dummy head was attached to the pendulum and orientated to face 0° angle of attack. The pendulum was raised up and measurement program was started. The program recorded the angle sensor data and both microphone signals simultaneously. Similar recording was done in every 30° intervals up to 180°.

The analysis of the recording was done by cutting a 328 ms recording around the maximum velocity and windowed DFT analysis was carried out.

6.2 Results

In this section the results obtained by the measurements explained in sections 6.1.2 to 6.1.5 are shown.

6.2.1 Effect of low-frequency tone to radiation of high frequency components from mouth in presence of flow

Figures 6.4 to 6.6 shows the measurement results for the low-frequency tone measurement described in section 6.1.2. In these figures the amplitude values of different tone frequencies are compared as follows. Measurement with flow is compared to static case, so negative values indicates that attenuation is due to flow or low-frequency tone. In every figure the flow directions of 0°, 90° and 180° are shown. In 6.4 only a single tone at 2853 Hz is presented in function of the amplitude of the low-frequency tone. In figure 6.5 a multi-tone with four frequency components is presented. Each of

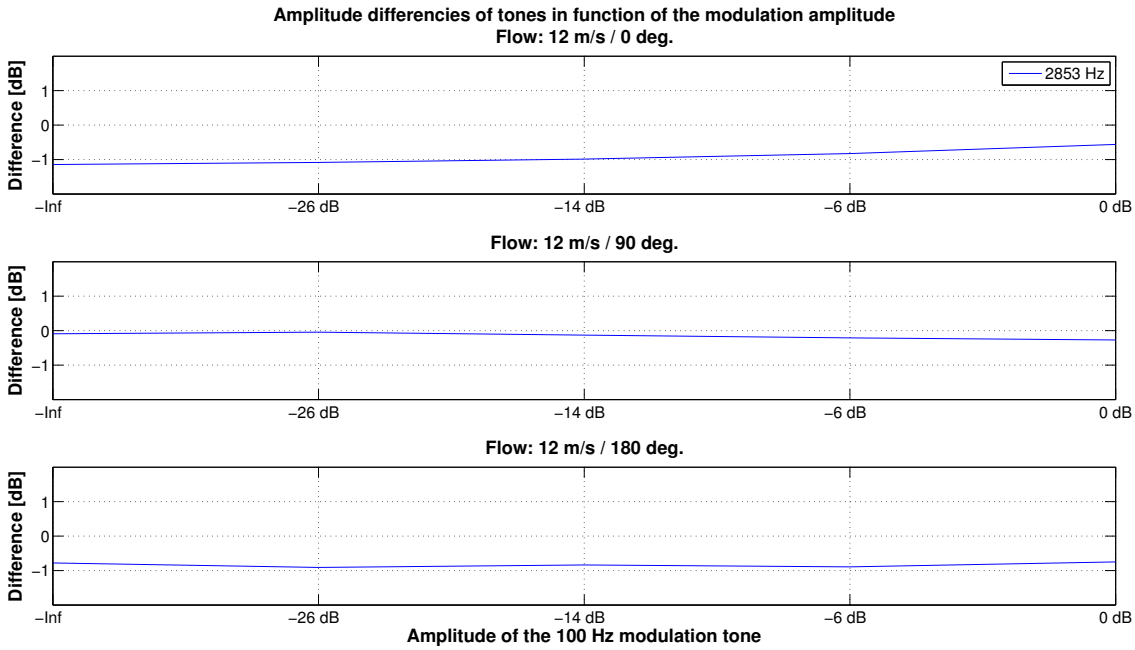


Figure 6.4: The effect of the amplitude of low-frequency tone to a single tone peak value.

the four frequency components are located at system resonance frequencies. In figure 6.6 a two tone signal is played. Both of the two tones are out of system resonance frequencies.

No effect were found from these measurement. This result indicates that the precence of low frequency tone does not introduce effects and the harmonic glottal excitation measurement signal can be replaced with non-harmonic multi-tone signal in next measurements.

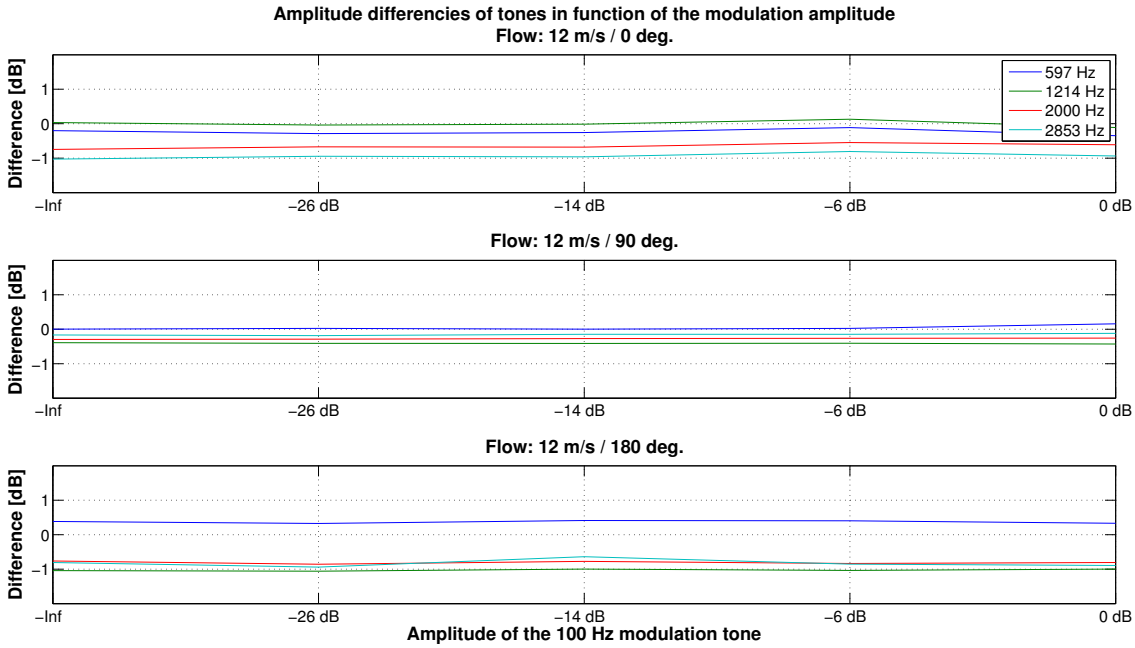


Figure 6.5: The effect of the amplitude of low-frequency tone to a tone values played at the system resonance frequencies.

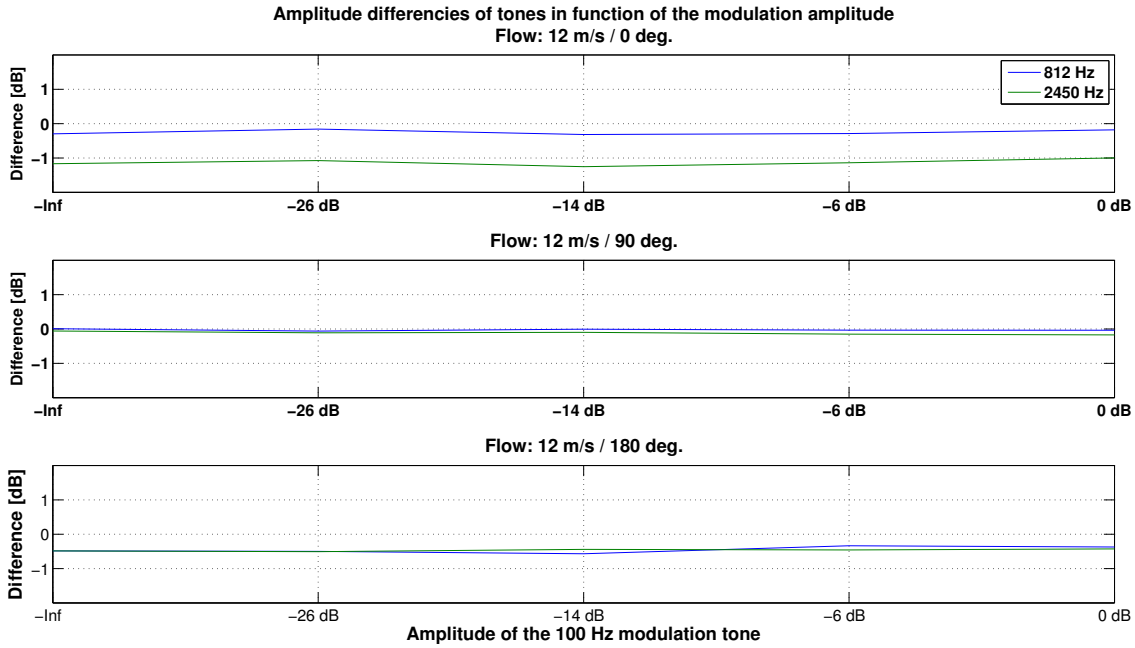


Figure 6.6: The effect of the amplitude of low-frequency tone values played at the antiresonance frequencies

6.2.2 Effect of flow to pressure inside vocal tract

This section shows the results of measurement on effect of flow to the pressure magnitude inside the vocal tract. The measurement procedure is introduced in section 6.1.3. Figures 6.7 to 6.13 presents the results for flow directions from 0° to 180° in 30° intervals. Figures are divided in two subfigures. In the upper subfigure, dashed lines are magnitude values of DFT of windowed measurement sequences. The solid lines are drawn by following the magnitude values for the frequency components of the excitation signal. The envelope of the DFT magnitude spectrum presents the magnitude response of the system within the frequency resolution of excitation signal. The solid envelope lines are raised 5 dB for better readability. Red lines present the data obtained by the static reference case and blue lines are shows measurement results with flow present. In left bottom of the upper figure, a flow direction related to head orientation is shown by an arrow. The lower subfigure shows the difference between the reference case and flow measurement.

From the measurement results it can be seen that the effects are minor and only a very few consistent effects can be pointed out. In all results, the low frequency fluctuation below the first formant frequency at 550 Hz is so random, that it is probably produced by the turbulent fluctuation close to mouth opening.

Within the frontal flow direction of 0°-30° (figures 6.7 and 6.8), a small dip occurs right above the second formant frequency around 1300 Hz. Similar effect below the formant is not visible, so it is hard to say whether the effect indicates the change in the bandwidth of the resonance or not. Similar but even weaker dampening can be seen within the third and fourth formants around 2100 Hz and 3000 Hz. Likewise in both of these directions, the higher frequency behavior seems to be the same. After a slightly attenuated fourth formant, there is a narrow band boost with magnitude of 3-4 dB below 4 kHz. Similar narrow band attenuation occurs below 6 kHz after which the response seems to be boosted.

For the flow direction of 60° (figure 6.9), a really small attenuation of the third formant frequency around 1900 Hz can be seen. The fluctuating below the first formant and between the first and second formants seem quite a random, so it is hard to say anything about them. Also the narrow band boost that occurred around 4 kHz for the flow directions of 0° and 30° has changed to a narrowband attenuation. With a 90° flow a small random fluctuation occurs around first two formants. The narrow band fluctuation around 4 kHz seen for frontal flow has now diminished or moved up a bit in frequency.

The for the 120°-150° flow directions the only effect around the four first formant frequencies seem to be a slight attenuation right above the second formant frequency around 1300 Hz. The same kind of behavior occurred with frontal flow directions of 0-30°. Likewise a narrowband boost below 4 kHz and attenuation around 6 kHz exist in 120° measurement like it existed in frontal directions. These narrow band high frequency fluctuations are not that clear for the 150° flow direction. The effect for the 180° flow direction seems to random and fluctuating. Also the raise of the

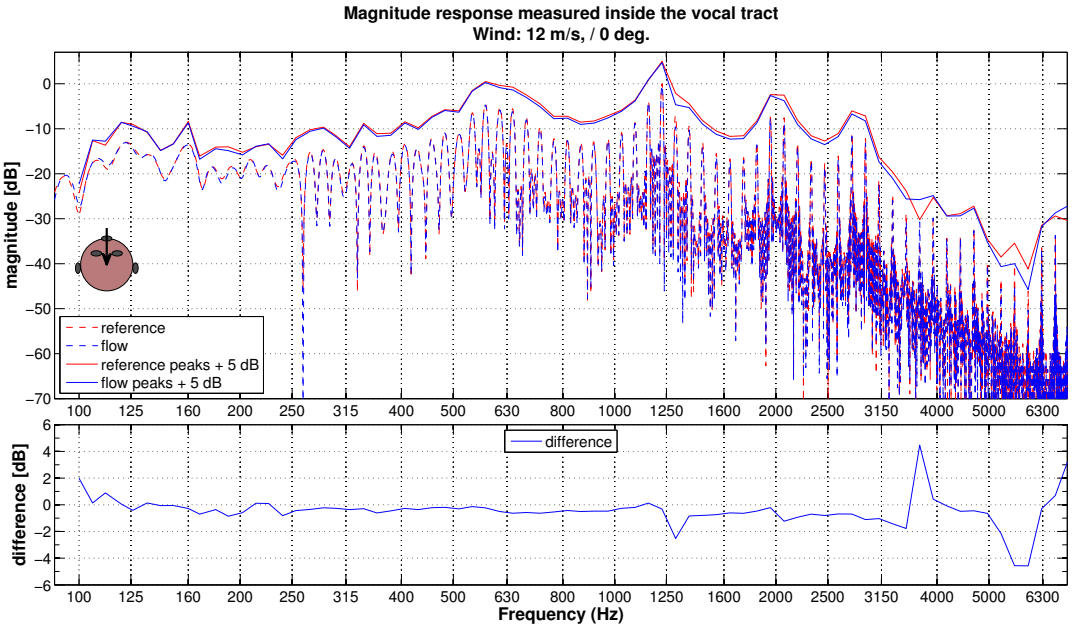


Figure 6.7: Magnitude response comparison inside the vocal track near mouth plane. The mouth is radiating in the upwind direction.

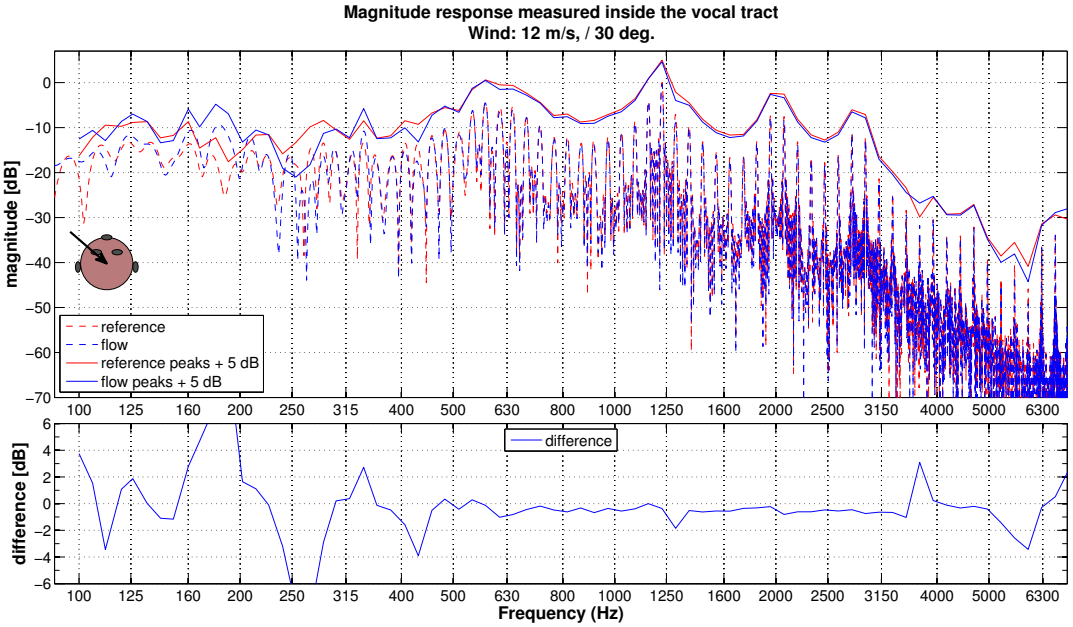


Figure 6.8: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 30°.

noise floor at higher frequencies indicates the disruption of the turbulence around the mouth opening, so it can be concluded that no clear effects are caused by the flow.

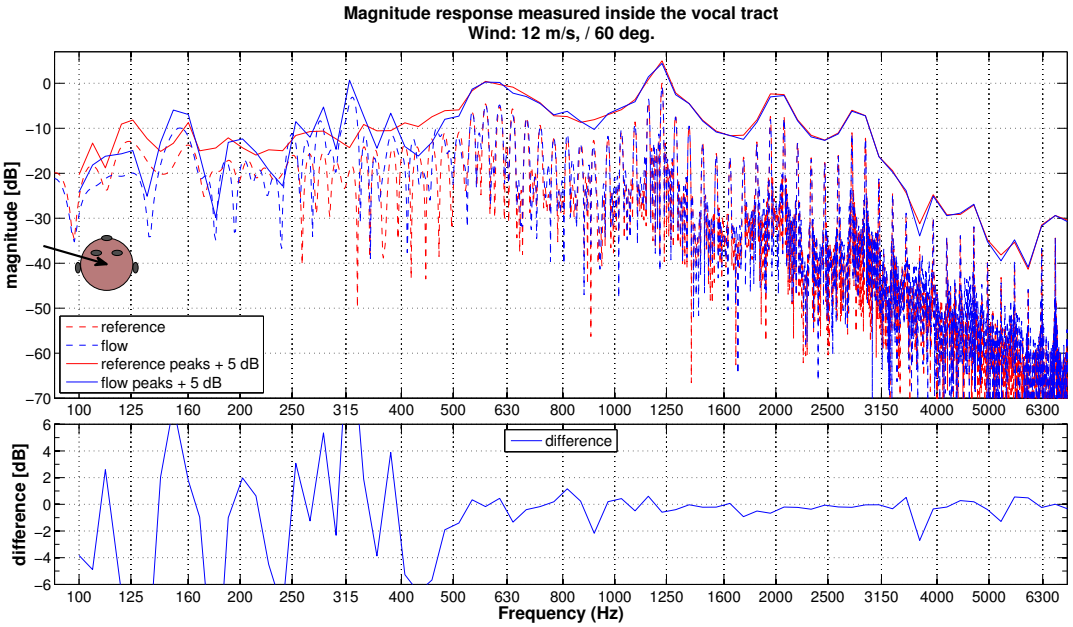


Figure 6.9: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 60°.

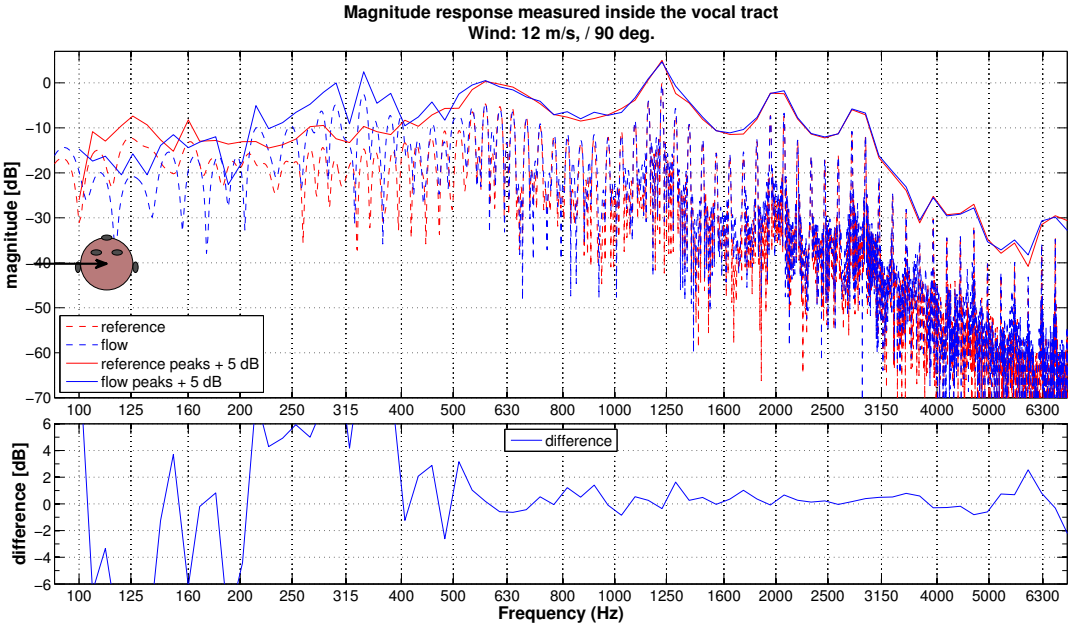


Figure 6.10: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 90°.

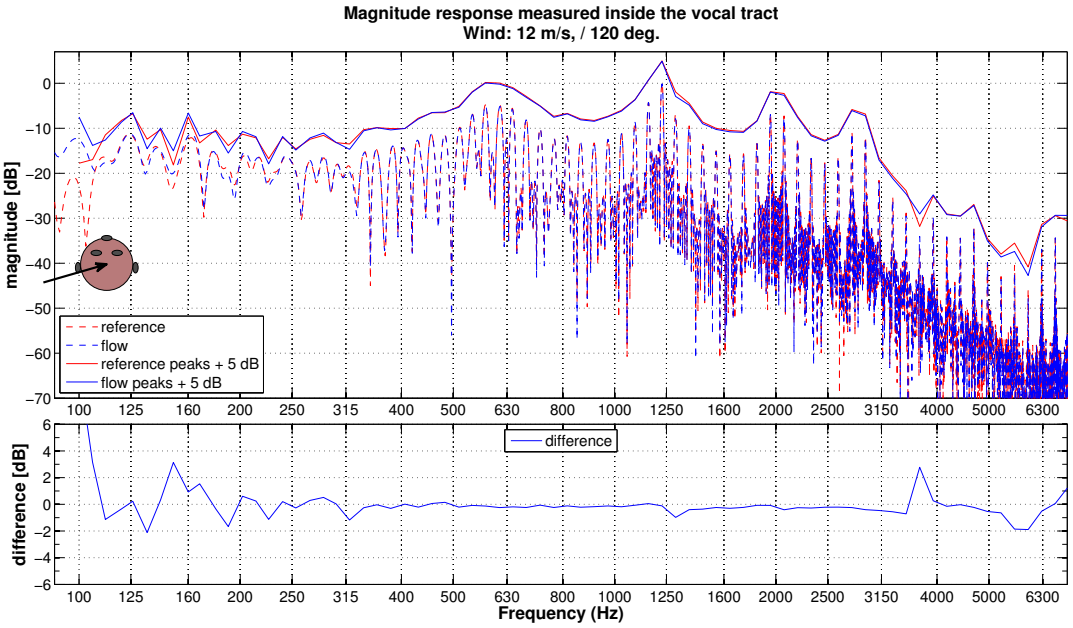


Figure 6.11: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 120°.

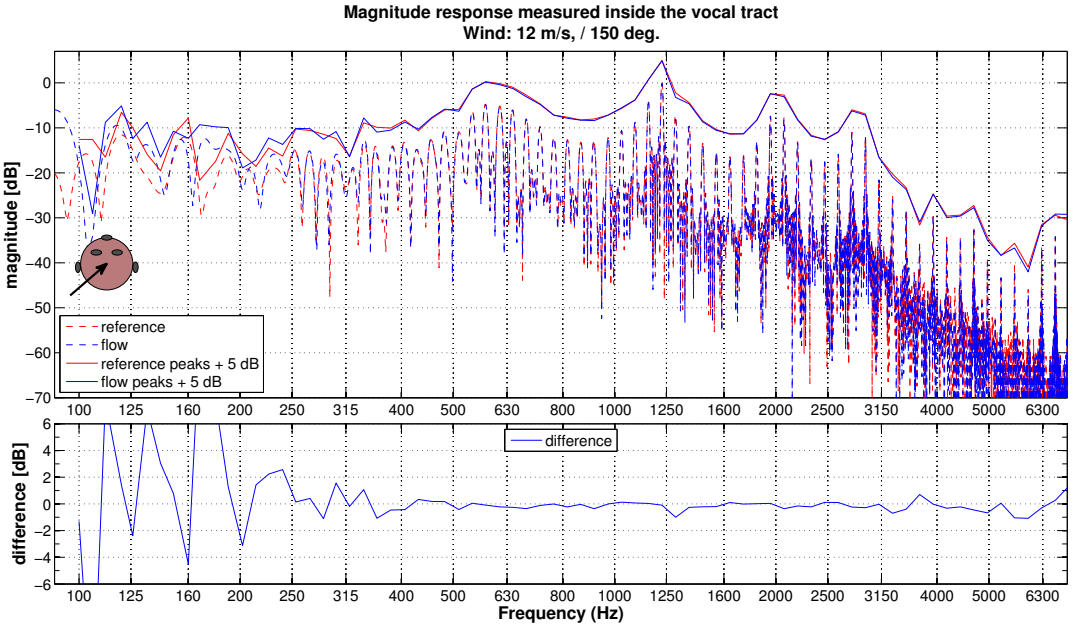


Figure 6.12: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 150°.

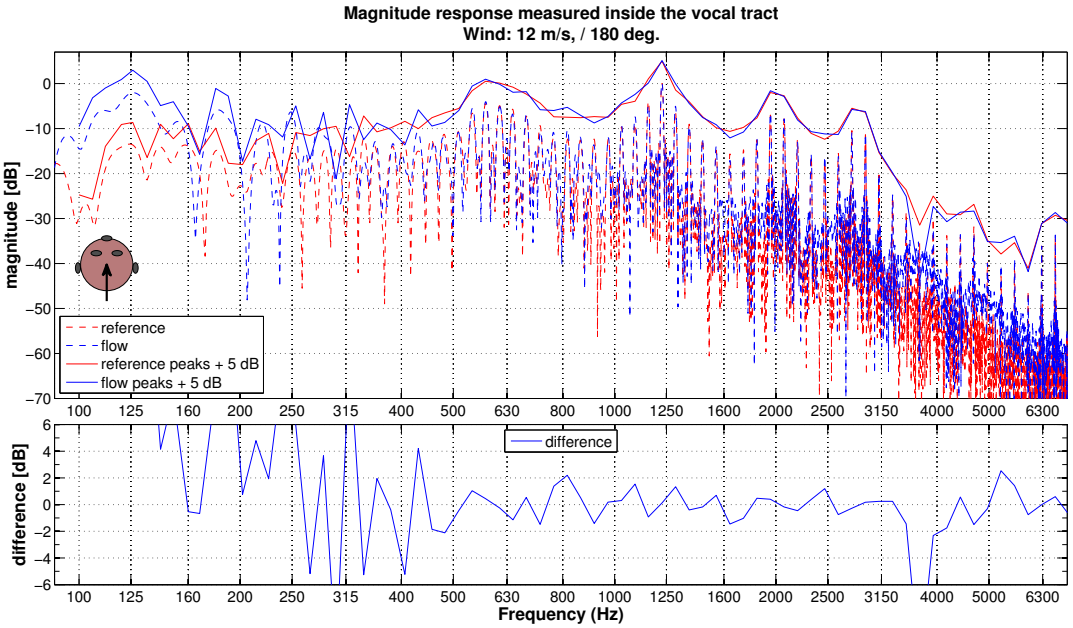


Figure 6.13: Magnitude response comparison inside the vocal track near mouth plane. The angle between the direction of radiation and upstream of the flow is 180°.

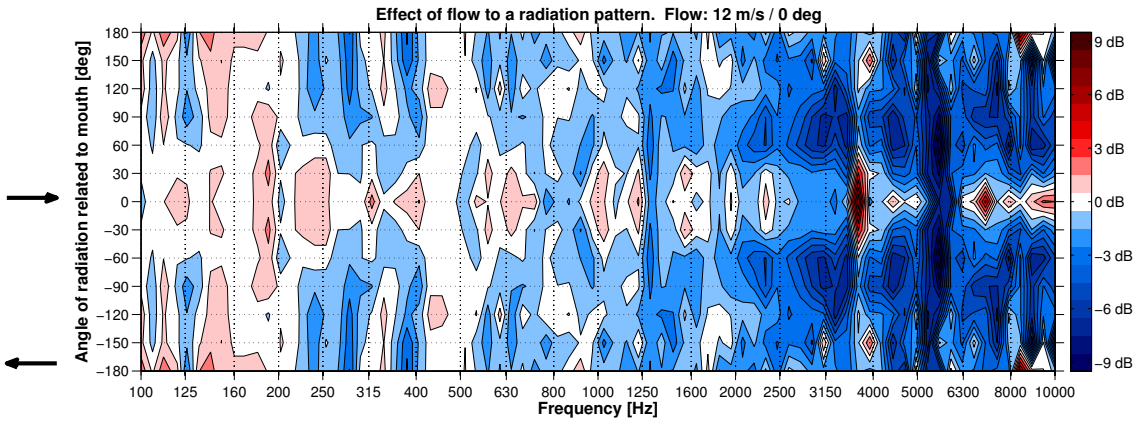


Figure 6.14: The change in the directional characteristics due the flow. The mouth is radiating in the direction of the upwind.

6.2.3 Effect of flow to directional characteristics of artificial speech source

In this section the results of polar response measurements of the artificial speech source are presented. Responses was measured in 30° intervals as described in section 6.1.4. The data is presented as a relative gain compared to reference case without a flow. The data in figures 6.14-6.20 are presented in contour plots, where a blue and red areas present attenuation and amplification, respectively. One colour presents 1 dB region around a integer value. The arrows on the left side of the plot - pointing to the right and left - shows the locations of the upstream and downstream stagnation points.

From the measurements a couple of consistent results can be pointed out. The most clear result seems to be the high frequency attenuation over 2 kHz. The attenuation seems to be pronounced in crosswind directions in every case, and the attenuation expands to even lower frequencies. Likewise a boost in the radiation seems to occur towards the upwind in every case, and the boost is most consistent and pronounced on frequencies below 2 kHz. The effects seems to be the same regardless of the relative direction of the mouth and the flow, altough a slight magnitude change in the effects can be seen.

6.2.4 Flow-induced noise at entrance of ear canals

The noise induced by flow was measured at the entrance of the ear canals of a dummy head, as described in section 6.1.5. The measurement was done with different flow directions to cover a semicircle around the head. The mean velocity at the 328 ms recording time was 11.9±0.1 m/s in measurements. Data in figures 6.21 and 6.22 shows a flow induced noise SPL at the ears on the upstream and downstream sides, respectively. One uniform colour pattern represents 5 dB region and every second

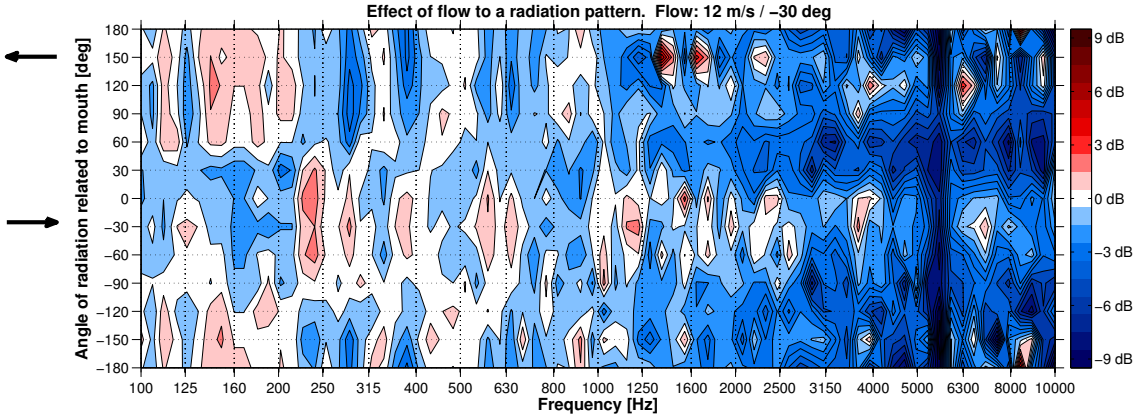


Figure 6.15: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 30°.

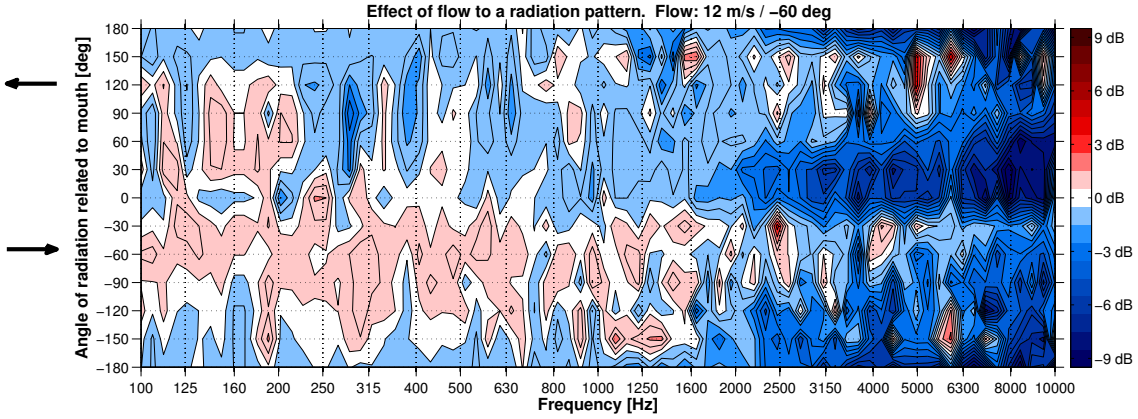


Figure 6.16: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 60°

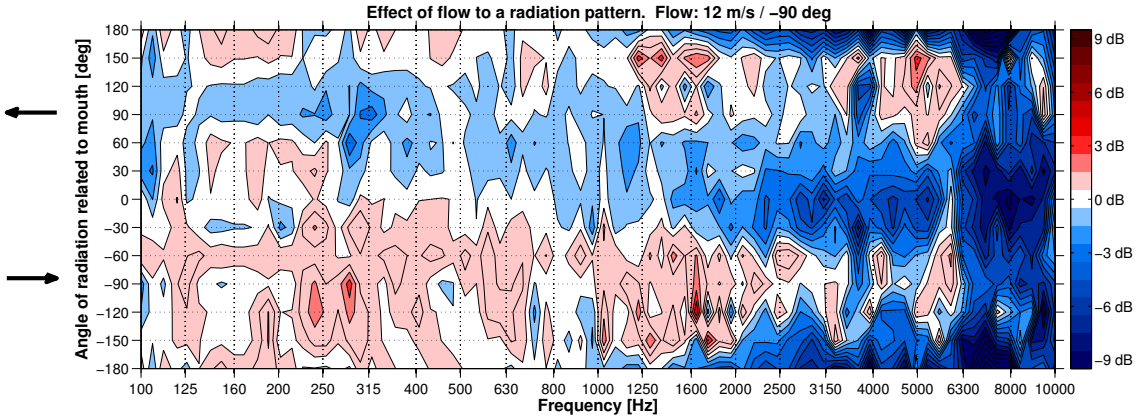


Figure 6.17: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 90°.

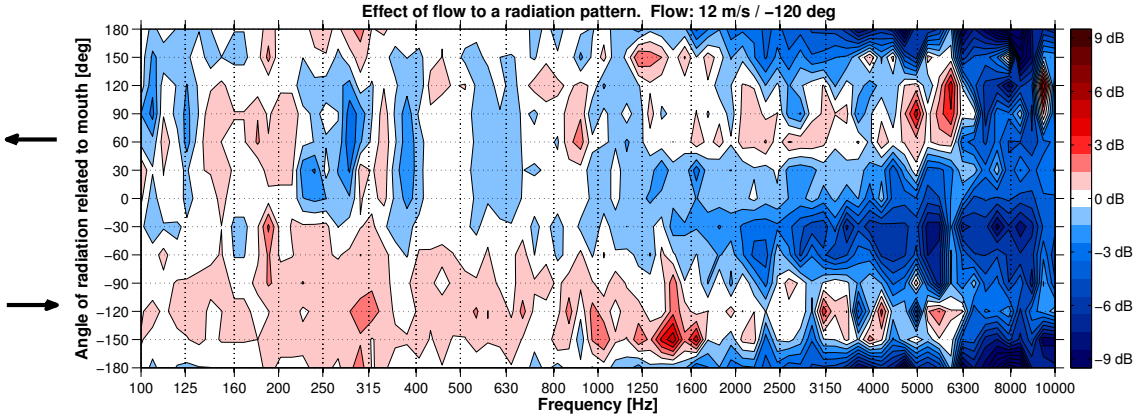


Figure 6.18: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 120°.

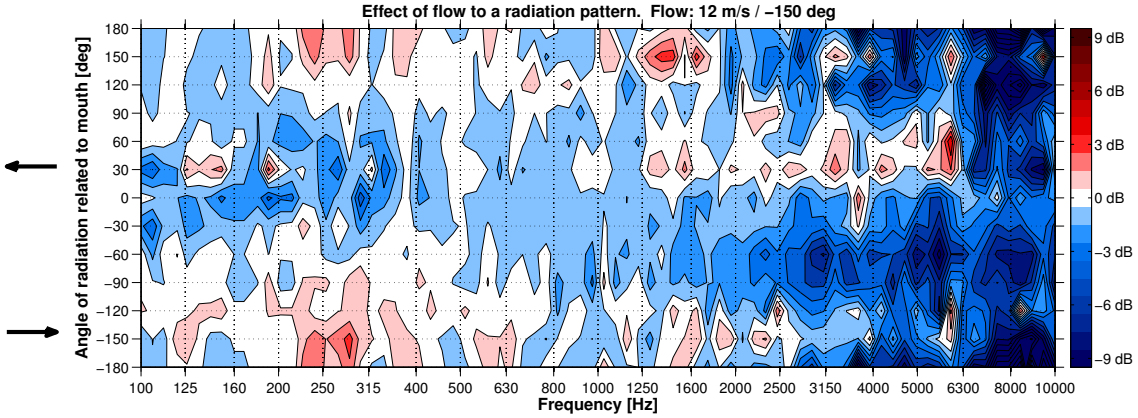


Figure 6.19: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 150°.

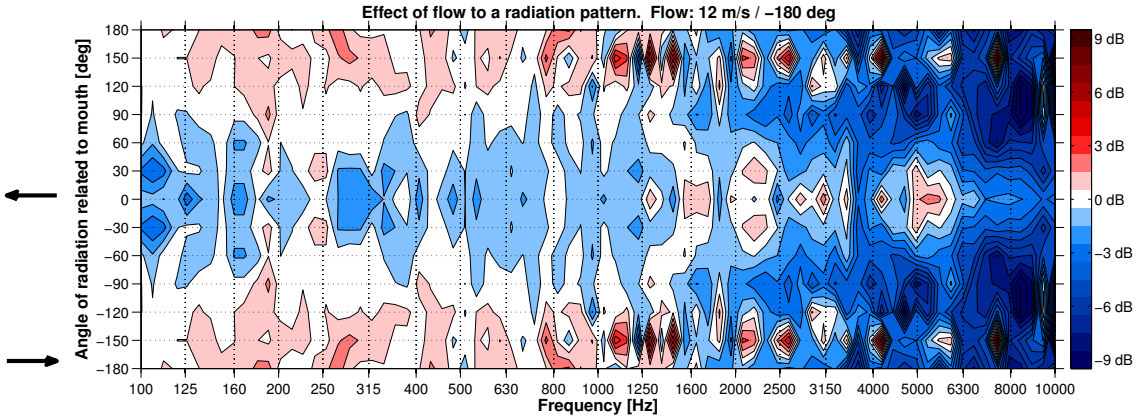


Figure 6.20: The change in the directional characteristics due the flow. The angle between the direction of radiation and upstream of the flow is 180°.

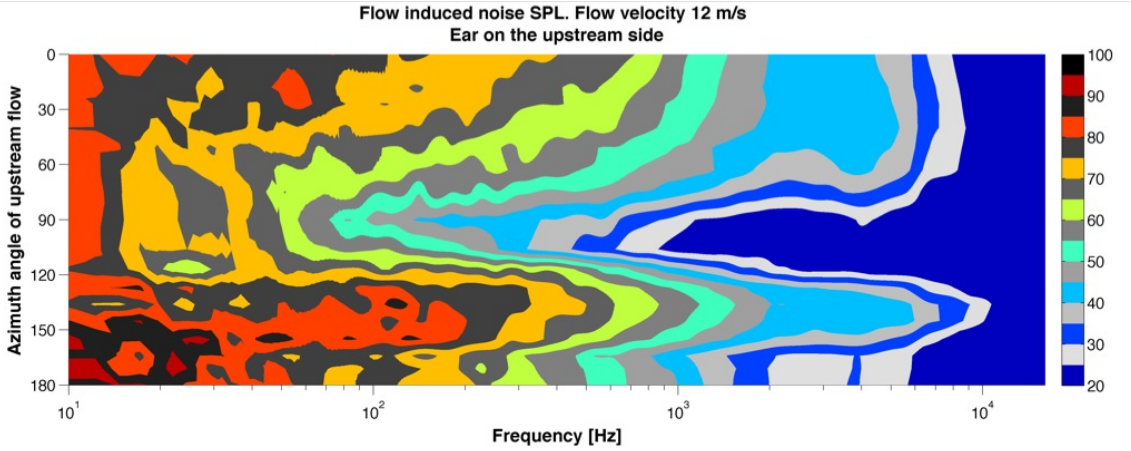


Figure 6.21: Flow induced noise SPL at the entrance of the ear canal on the upwind side. The constant color presents 5 dB region.

5 dB region is plotted as a grayscale to help identifying different regions. Figure 6.23 shows a level differences between the ears.

From these results it can be seen that the noise has low pass characteristics and can exceed 100 dB SPL below 100 Hz. The slope and shape of the low pass characteristics changes in function of the flow direction. The spectrum differs between the ears a lot for other than front and back flow directions. The results show that when the flow directions starts to move from the front to the side, a noise levels decreases at the upwind side and the low frequency levels raises at the down wind side. It can be noted that with direct side wind the simultaneous levels at the both ears are lowest and with around 120-150° the simultaneous levels on both ears are highest. Still the highest levels occurs at the down wind side for around 30-60° flow direction. The most important result for this measurement can be seen from the figure 6.23. It shows that a level differences between the two ears can be huge with different flow directions. The most pronouced effects are around 60° and 120° directions, where a low frequency levels can be about 30 dB higher on the down side while at the same time the high frequencies can be up to 25 dB quieter at the down wind side. This can be seen a a really significant change in the shape of the spectrum of flow-induced noise. More detailed graphs with each measured flow directions are shown in appendix A.2.

6.3 Evaluation of measurement method

The measurements done during the work of this thesis proved that the pendulum system was capable of producing a reproducible motion, whose velocity corresponded to wind of a strong breeze. The most important factor for this measurement setup was it's passive motion generation method. It provided minimal background noise levels due to a flow generation itself. Another important feature was anechoic conditions.

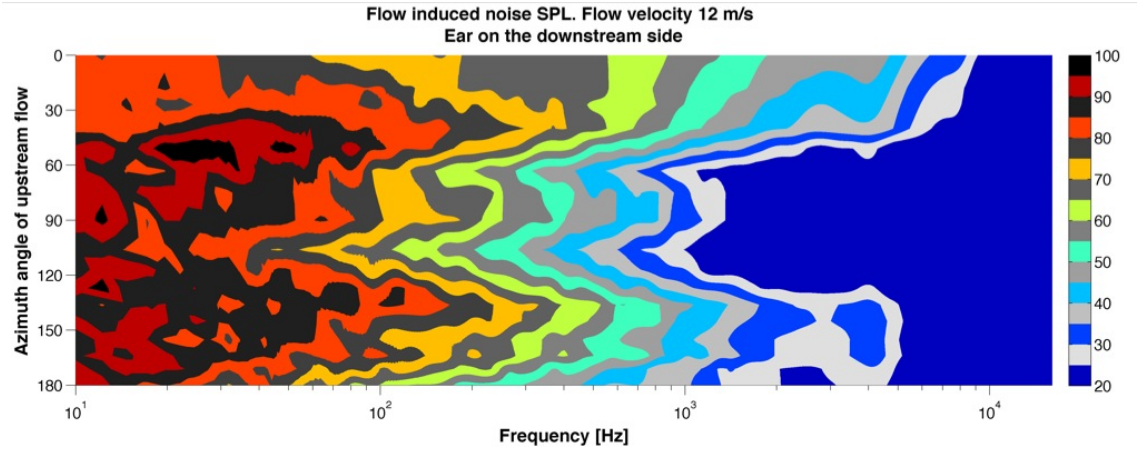


Figure 6.22: Flow induced noise SPL at the entrance of the ear canal on the downwind side. The constant color presents 5 dB region.

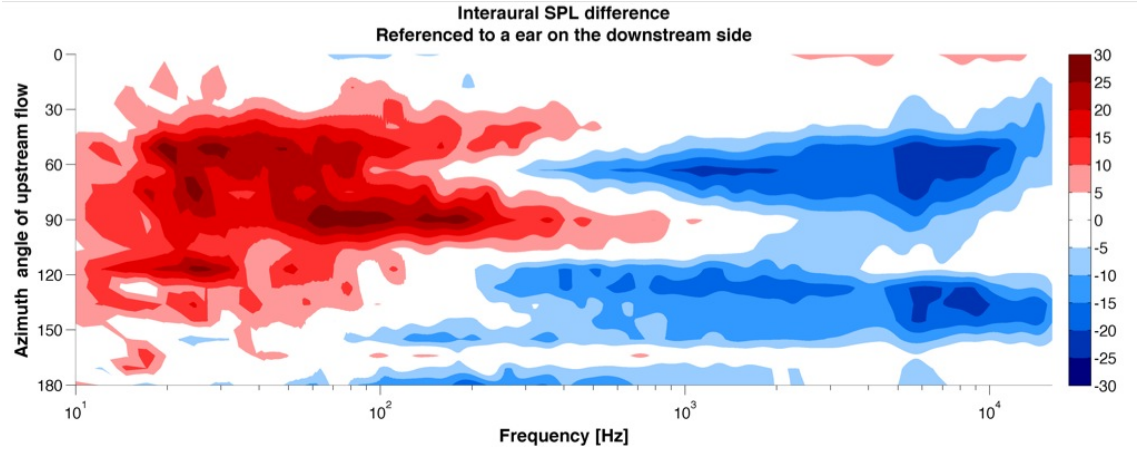


Figure 6.23: Interaural level difference on the flow induced noise in 5 dB steps. Difference is referenced to a pressure at the ear on the downwind side. White region indicates a $\pm$ 5 dB region around 0 dB.

The noise levels inside the anechoic chamber or the noise generated by the pendulum frame, bearing or other parts didn't incur problems at any point. The flow induced noise due to the moving objects was higher than was first expected. This enforced the development of the artificial speech source and excitation signals to overcome the noise levels due to flow. Partly because of the high flow noise levels, the perspective of this thesis expanded to cover the flow-induced noise levels in ear canals.

The pendulum structure remained consistent, stable and rigid trough all measurement. The setup and rigid guying of the pendulum frame in to the anechoic chamber was two days work. The pendulum assembly was possible to handle by one man. Portability enabled the deallocation of the anechoic room for other studies after the measurement period.

Although the pendulum served very well in its task, there are multiple things that had to be taken into account and thus reduced the flexibility of the system. The major issue of the system and measurement procedure was a short time window for data acquisition. The issue originates from the acceleration of the pendulum, and thus the changing nature and velocity of the device under test. Despite of that, the velocity remained quite stable over the 100 ms period around the vertical 0°angle. The usable time window during one swing was still quite small. The small time window wasn't that prominent issue in measurements where the microphones moved along the pendulum. Even more notable drawback due to moving nature of the source can be seen in the directivity measurement. During the analysis window, the source moved one meter (section 6.1.4). During that time the angles to microphones on static hemispherical array changed also. The change of angle was largest in nominal 90° direction, being $\pm 75 - 105°$.

Without the detailed analytical analysis or inspection of the problem trough computational fluid dymamics (CFD) tools - that is out of scope of this thesis - the errors or inaccuracy due to angle shift and doppler removal time mapping processing remains unclear.

Chapter 7

Discussion

The performance decrease due to the flow induced noise on microphones in different applications has been in interest as long as microphones has been produced and used in outdoor conditions. A flow generator for microphone and windshield performance measurements has been standardized and widely used by microphone industry. On the another end, a theory for sound propagation trough atmosphere has been used with good correlation to experimental data and measurement. The area of acoustics related to moving medium within the distances of human speech communication has not been widely studied. This thesis took some insight into this area by introducing phenomena related to this problem together with measurements of the effects of flow to speech communication.

The pendulum system with a static microphone array overtook most of the problems with the price of some complexity in the post processing. Reasonable flow noise levels were obtained with a measurement technic where a microphone was mounted on the pendulum frame as described in 6.1.2. A lower velocity at the microphone mounting position suppressed the wind noise problems.

The measured results relad to sound radiation and transmission cannot be compared with any other case, as no other study on the subject was found. More repetetitions should have been made in order to get some statistical data of the results. A real world observation of the possible phenomena would help in focusing the study in more specified area. One interesting subject would be to study the effect of the flow on function of the glottis. If a considerable increase of the subraglottal pressure would exist, how muct more work would be needed to maintain the transglottal pressure and vocal cord action compared to normal non-windy conditions.

Chapter 8

Conclusions

A natural wind is complex phenomena and the movement of medium is constantly changing in the form of turbulences with different scales. The flow generation over large areas is challenging task. During the work of this thesis the pendulum system for producing a relative movement between the sound source and medium was built. The relative movement produced by pendulum could be seen as laminar flow over the source in rest. With the setup described in this work, a set of measurements was carried out in order to study the effects of flow on speech production, transmission and perception on speech communication.

The measurements were laborious to carry out and lot of data was recorded from different combinations. Within the measurement results obtained, it can been said that the effect of flow to speech production and function of the vocal tract are minor. More clear results can be seen in far field radiation, where high frequency attenuation in all directions takes place. The results show also that almost regardless of the mouth orientation, increased gain of radiation takes place in the upwind direction on the frequency range of first formants. The most significant results of this thesis can be seen in the flow induced noise levels on ears. Especially the attention was paid to the differences on binaural spectral balance of the noise.

Considering the introduced literature on known effects of the wind on sound transmission and the measurement data of this thesis, it can be concluded that the most prominent effect of disturbing the human speech communication in windy conditions is the flow-induced noise in ear canals. Wind induced shadow zones do not occur within a distance of a shout, but flow noise levels can be quite intense around the shouting person and the listener.

Bibliography

[1] P. Alku. Glottal wave analysis with Pitch Synchronous Iterative Adaptive Inverse Filtering. *Speech Communication*, 11(2-3):109–118, Jun. 1992.

[2] P. Alku, H. Tiitinen, and R. Näätänen. A method for generating natural-sounding speech stimuli for cognitive brain research. *J. Clin. Neurophysiol.*, 110:1329–1333, Mar. 1999.

[3] J. Anderson. *Fundamentals of aerodynamics*. McGraw-Hill series in aeronautical and aerospace engineering. McGraw-Hill, 2001.

[4] J. Awrejcewicz. *Classical Mechanics: Kinematics and Statics*. Advances in mechanics and mathematics. Springer New York, 2012.

[5] G. Baker and J. Blackburn. *The Pendulum: A Case Study in Physics*. OUP Oxford, 2005.

[6] J. Benesty and Y. Huang. *Springer Handbook of Speech Processing*. Springer handbooks. Springer, 2008.

[7] J. Blauert. *Spatial hearing: the psychophysics of human sound localization*. Mit Press, 1997.

[8] J. C. Bleazey. Experimental determination of the effectiveness of microphone wind screens. *J. Audio Eng. Soc*, 9(1):48–54, 1961.

[9] S. Braun. Time-domain formulation of the doppler effect. *The Journal of the Acoustical Society of America*, 59(6):1495–1497, 1976.

[10] E. B. Brixen. Microphones, high wind and rain. In *Audio Engineering Society Convention 119*, 10 2005.

[11] M. Brock. Wind and turbulence noise of turbulence screen, nose cone, and sound intensity probe with wind screen. Brüel & kjær technical review no. 4, Aalto SCI, 1986.

[12] M. J. Crocker. *Handbook of Acoustics*. A Wiley-Interscience Publication. Wiley, 1998.

[13] F. J. Fahy. *Foundations of Engineering Acoustics*. Academic Press, 2001.

[14] M. Fishburn, G. Colin-Thome', and P. Alway. Generating a "silent" flow of air to measure the aerodynamic and acoustical performance of a louvre under

acoustical laboratory conditions. In *14th International Congress on Sound & Vibration*, 7 2007.

[15] J. Flanagan. *Speech analysis; synthesis and perception*. Kommunikation und Kybernetik in Einzeldarstellungen. Springer-Verlag, 1972.

[16] J. L. Flanagan. Analog measurements of sound radiation from the mouth. *The Journal of the Acoustical Society of America*, 32(12):1613–1620, 1960.

[17] M. Hiipakka, M. Tikander, and M. Karjalainen. Modeling the external ear acoustics for insert headphone usage. *J. Audio Eng. Soc*, 58(4):269–281, 2010.

[18] J. Hill, J. Blotter, and T. Leishman. Harsh environment windscreen analysis and design. *The Journal of the Acoustical Society of America*, 116(4):2618–2618, 2004.

[19] P. Hofman and A. Van Opstal. Binaural weighting of pinna cues in human sound localization. *Experimental Brain Research*, 148(4):458–470, 2003.

[20] J. Huopaniemi, K. Kettunen, and J. Rahkonen. Measurement and modeling techniques for directional sound radiation from the mouth. In *Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop on*, pages 183–186, 1999.

[21] IEC60268. Sound system equipment - part 4: Microphones, edition 3, 2004.

[22] U. Ingard. A review of the influence of meteorological conditions on sound propagation. *The Journal of the Acoustical Society of America*, 25(3):405–411, 1953.

[23] Microflown. *The Microflown E-Book, Chapter 15*, chapter 15. Microflown, 2009.

[24] H. Olson. *Acoustical Engineering*. Van Nostrand, 1957.

[25] T. Parsons. *Voice and speech processing*. MCGRAW HILL SERIES IN ELECTRICAL AND COMPUTER ENGINEERING. McGraw-Hill, 1987.

[26] N. Pedersen and O. Sø rensen. The compound pendulum in intermediate laboratories and demonstrations. *American Journal of Physics*, 45(10):994–998, 1977.

[27] F. Petrescu. *A New Doppler Effect*. Books on Demand, 2012.

[28] T. Rossing. *Springer Handbook of Acoustics*. Springer handbooks. Springer, 2007.

[29] M. Schneider. Wind & weather. In *Audio Engineering Society Convention 116*, 5 2004.

[30] H. Smith. *Illustrated Guide to Aerodynamics*. TAB BOOKS Inc., 1985.

[31] K. Sugiyama and H. Irii. Comparison of the sound pressure radiation from a

prolate spheroid and the human mouth. *Acta Acustica united with Acustica*, 73(5):271–276, May.

[32] M. Takanen, T. Raitio, O. Santala, P. Alku, and V. Pulkki. Fusion of spatially separated vowel formant cues. *The Journal of the Acoustical Society of America*, 134(6):4508–4517, 2013.

[33] G. Von Békésy. *Experiments in Hearing*. McGraw-Hill series in Psychology. McGraw, 1960.

[34] J. A. Zakis. Wind noise at microphones within and across hearing aids at wind speeds below and above microphone saturation. *The Journal of the Acoustical Society of America*, 129(6):3897–3907, 2011.

[35] E. Zwicker and H. Fastl. *Psychoacoustics: facts and models*. Springer series in information sciences. Springer-Verlag, 1990.

Appendix A

Appendix A

A.1 Directivity plots

This chapter presents polar plot figures in third octave bands. In average four tone from the multi-tone signal is included in a one third-octave band. Averare of these tone components are calculated and presented. Figures A.1 to A.20 compares the effect of direction within a one band at a time. In all polar plots red line is the reference case without the flow and blue line presents the response with 12 m/s flow. Magenta coloured line point to the direction of upwind direction. The angle labels are related to the mouth direction and spacing of a concentrical lines corresponds 6 dB difference in magnitude.

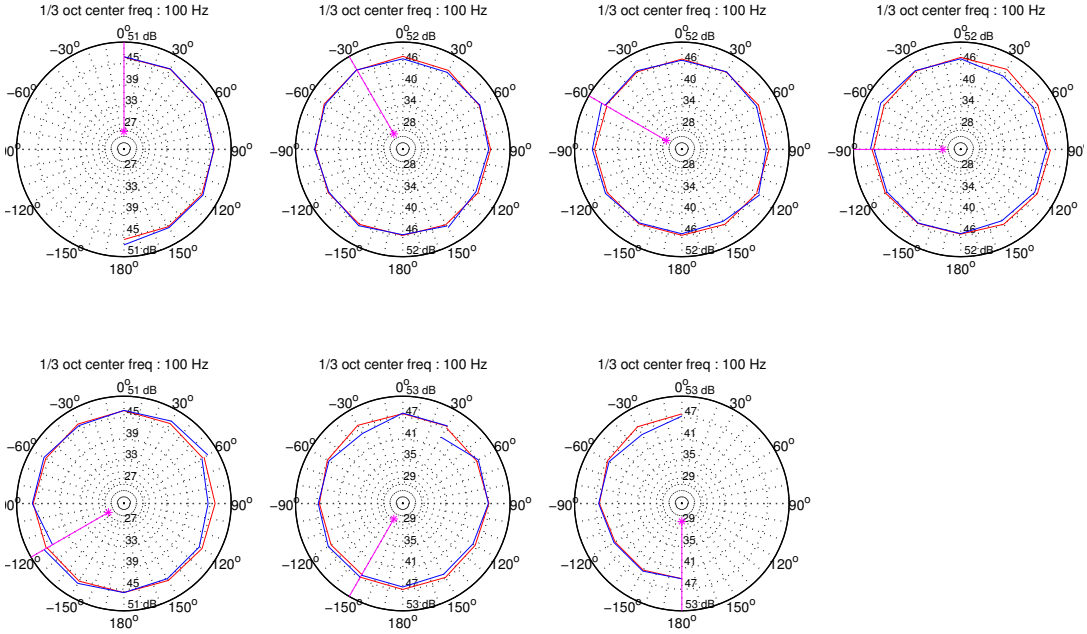


Figure A.1: Directivity comparison, 100 Hz band.

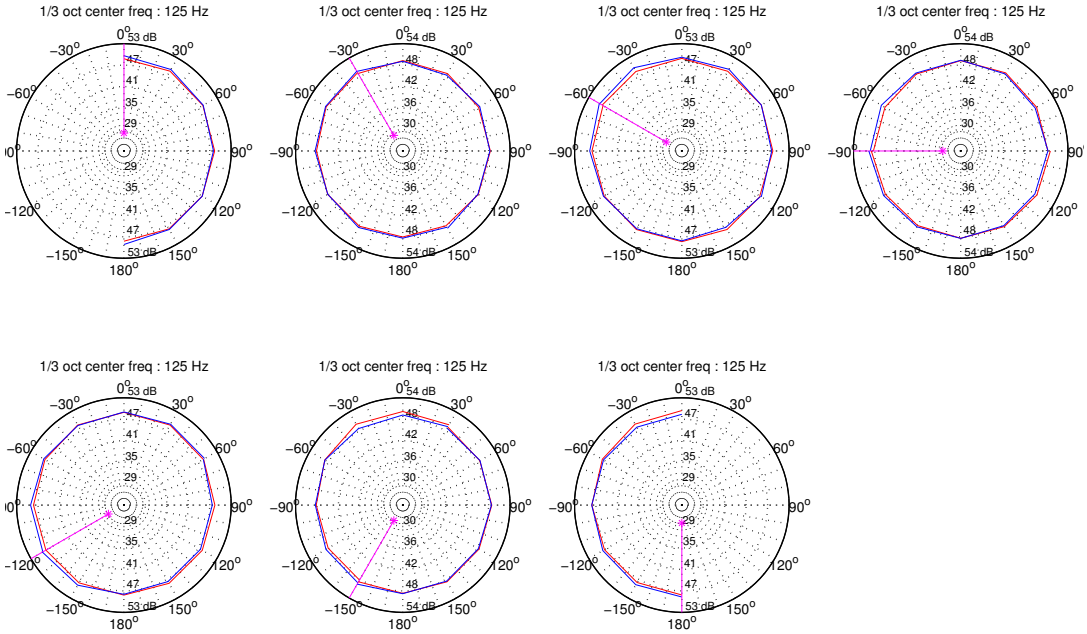


Figure A.2: Directivity comparison, 125 Hz band.

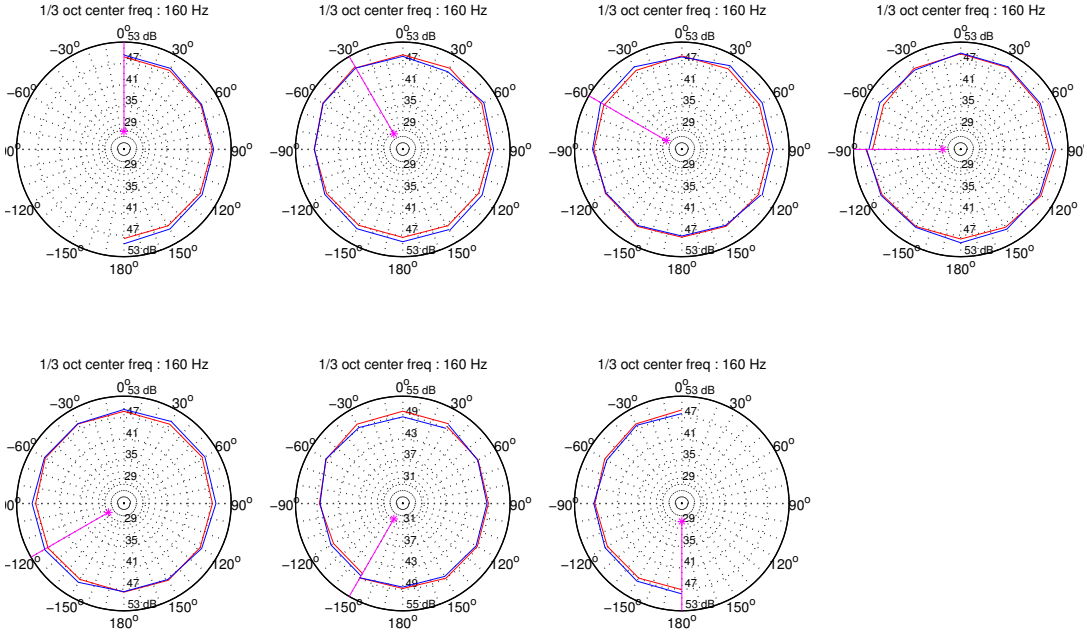


Figure A.3: Directivity comparison, 160 Hz band.

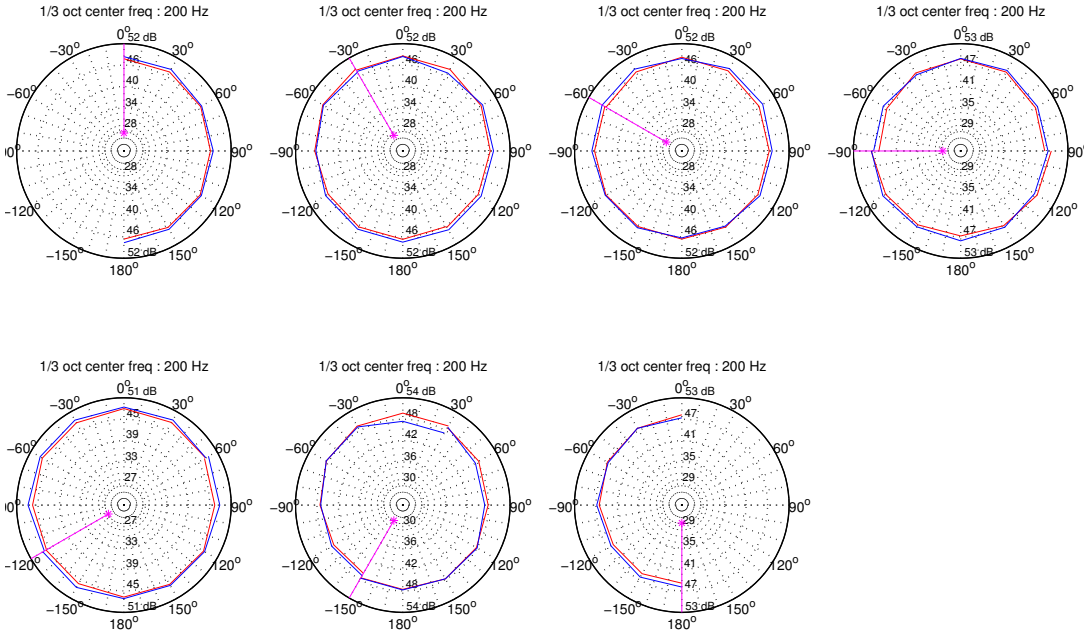


Figure A.4: Directivity comparison, 200 Hz band.

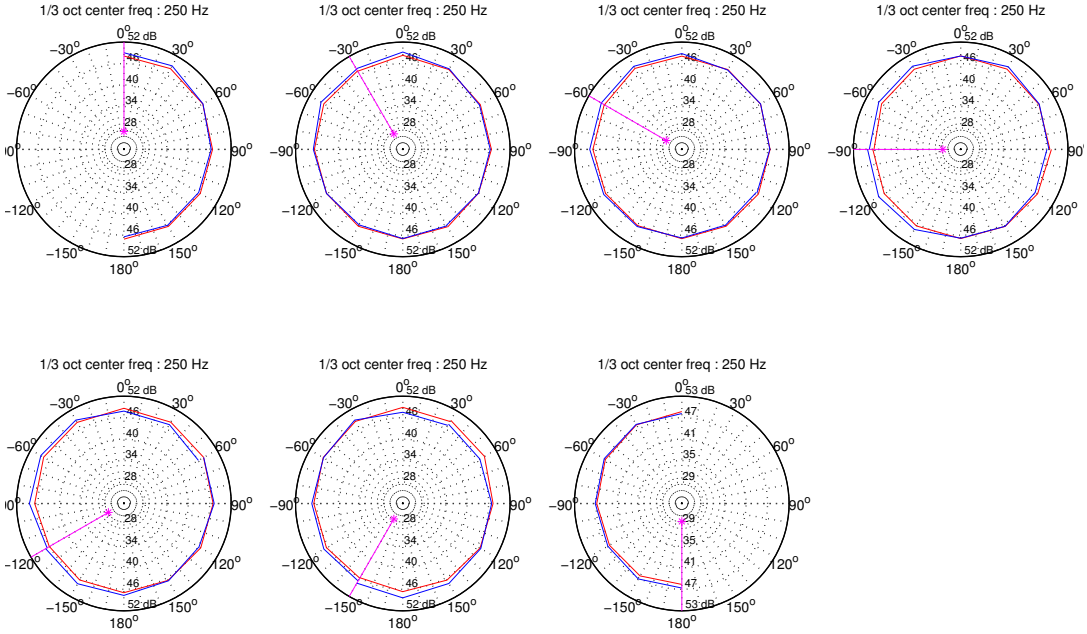


Figure A.5: Directivity comparison, 250 Hz band.

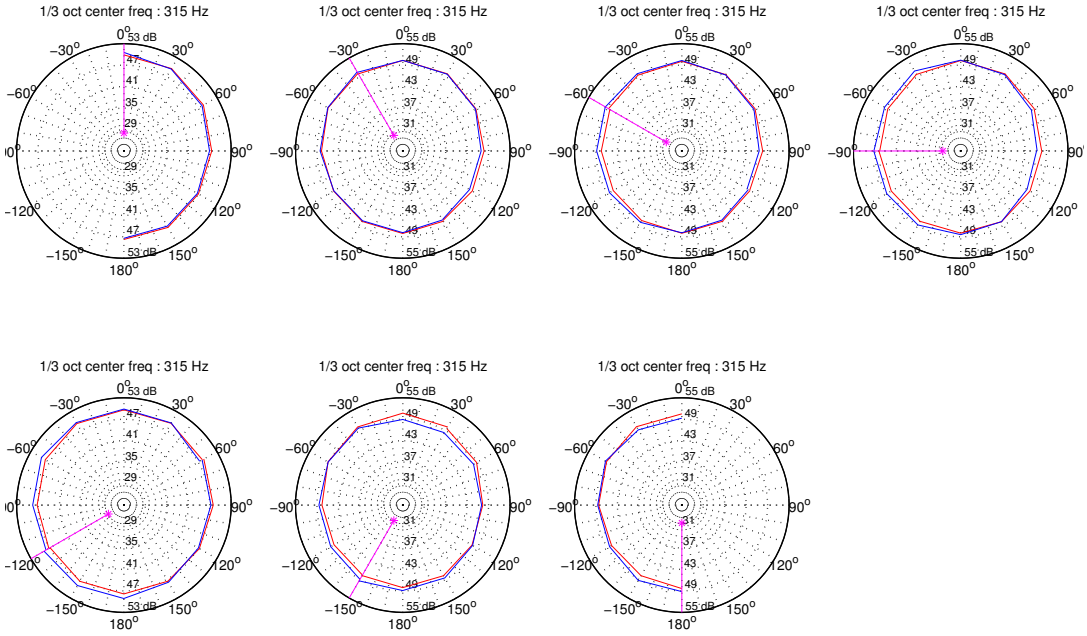


Figure A.6: Directivity comparison, 315 Hz band.

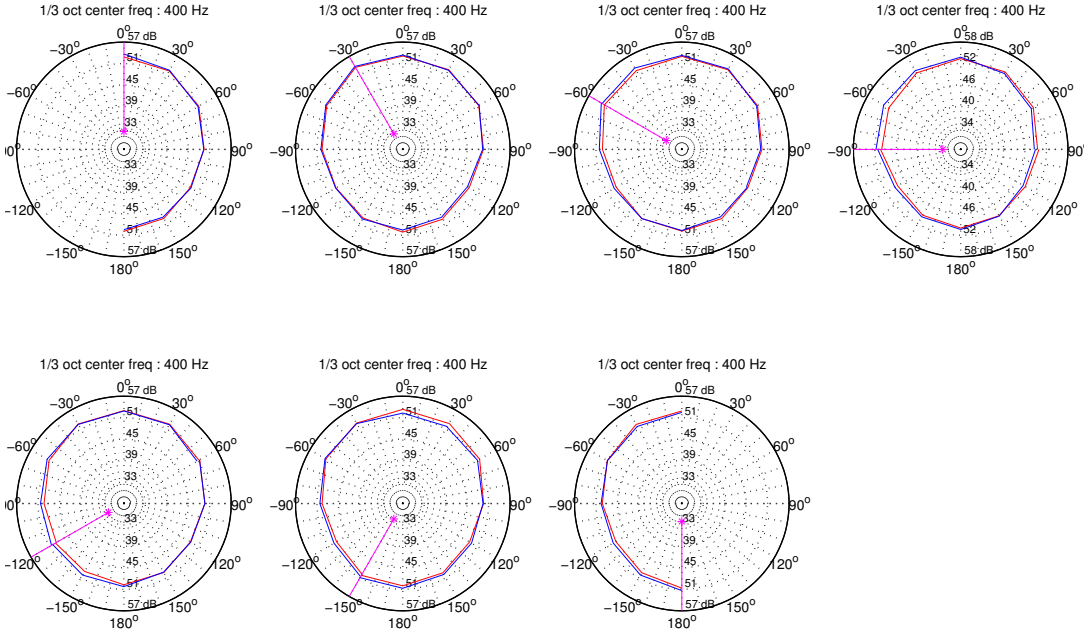


Figure A.7: Directivity comparison, 400 Hz band.

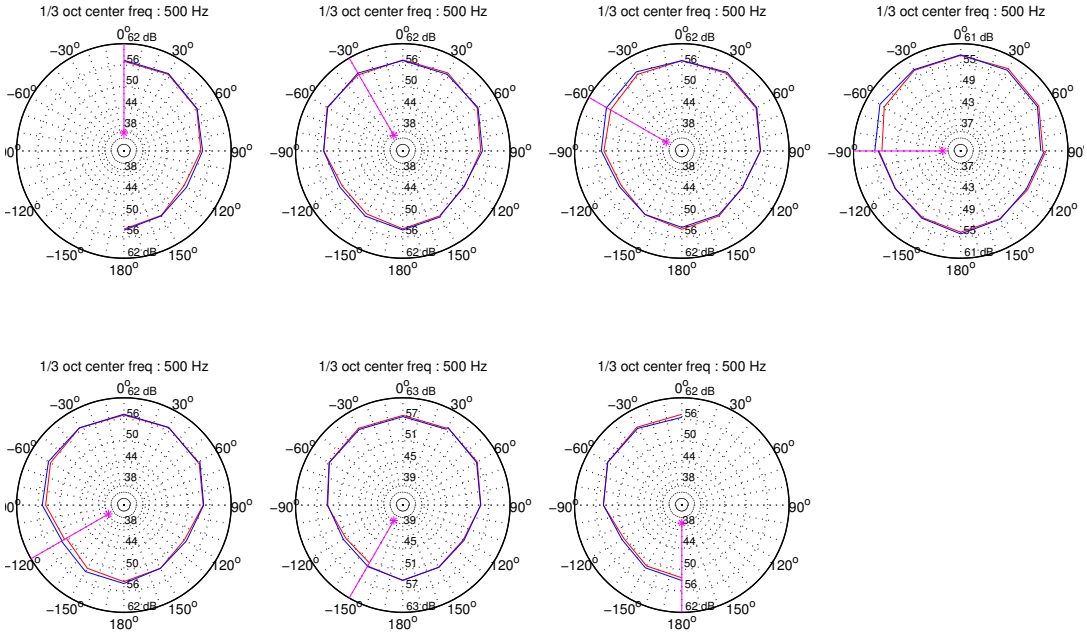


Figure A.8: Directivity comparison, 500 Hz band.

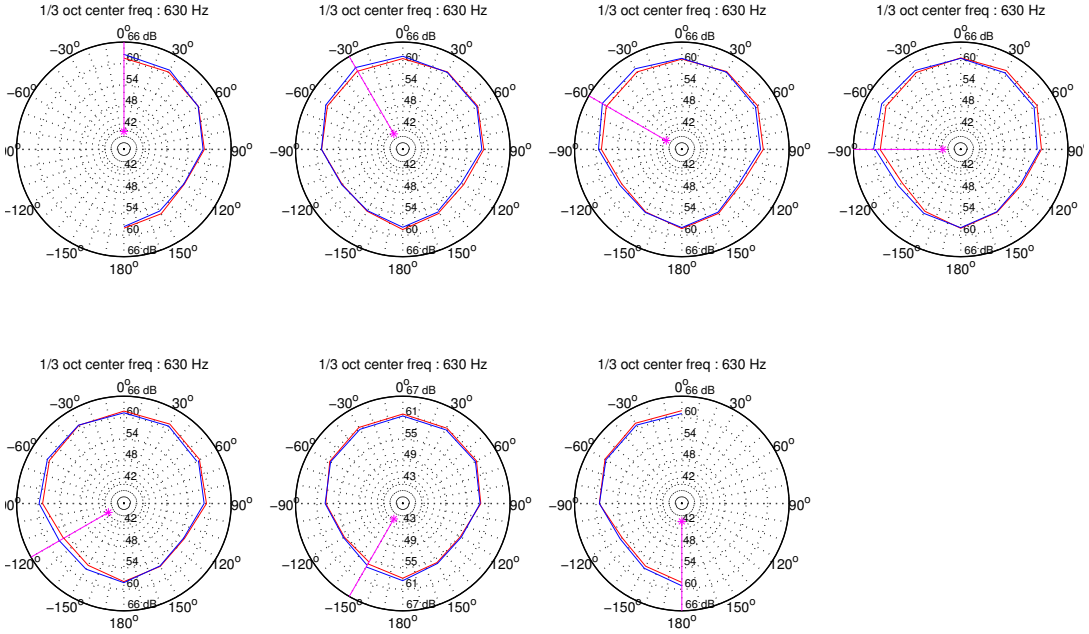


Figure A.9: Directivity comparison, 630 Hz band.

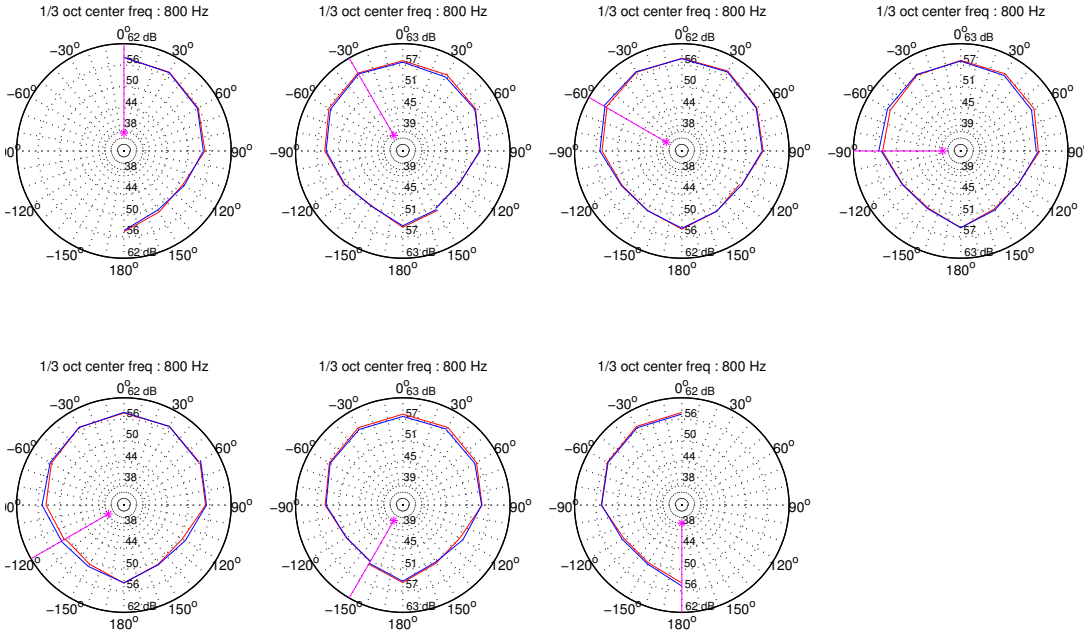


Figure A.10: Directivity comparison, 800 Hz band.

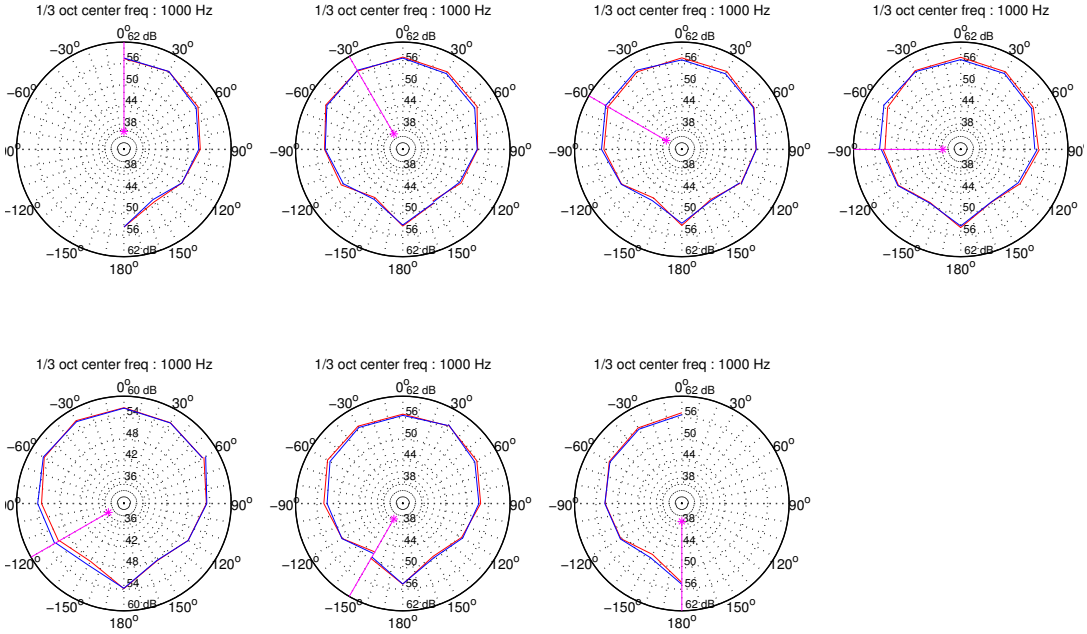


Figure A.11: Directivity comparison, 1000 Hz band.

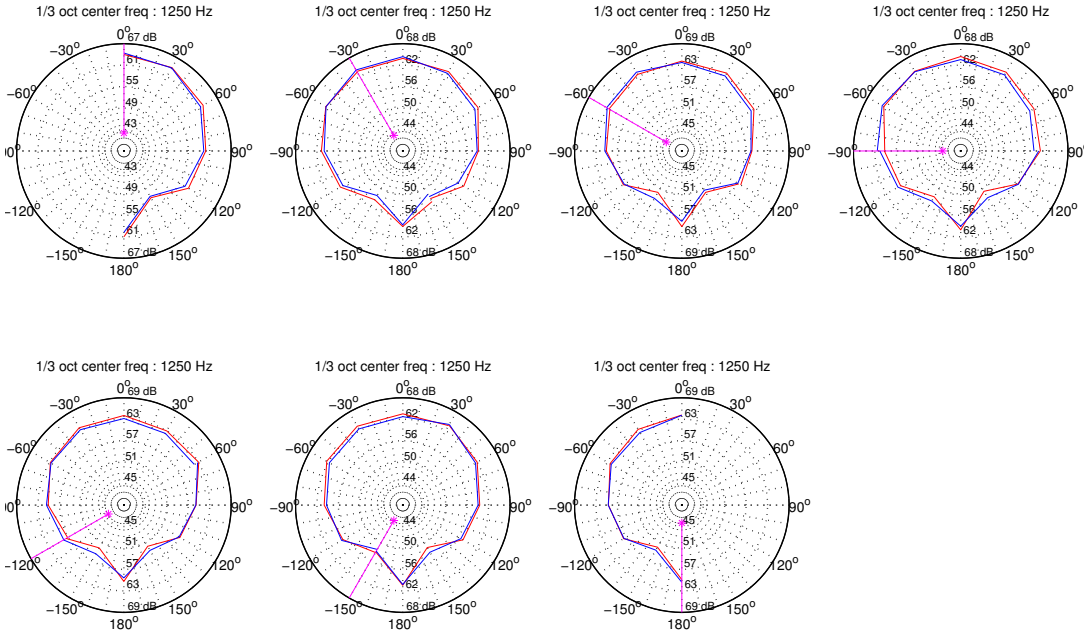


Figure A.12: Directivity comparison, 1250 Hz band.

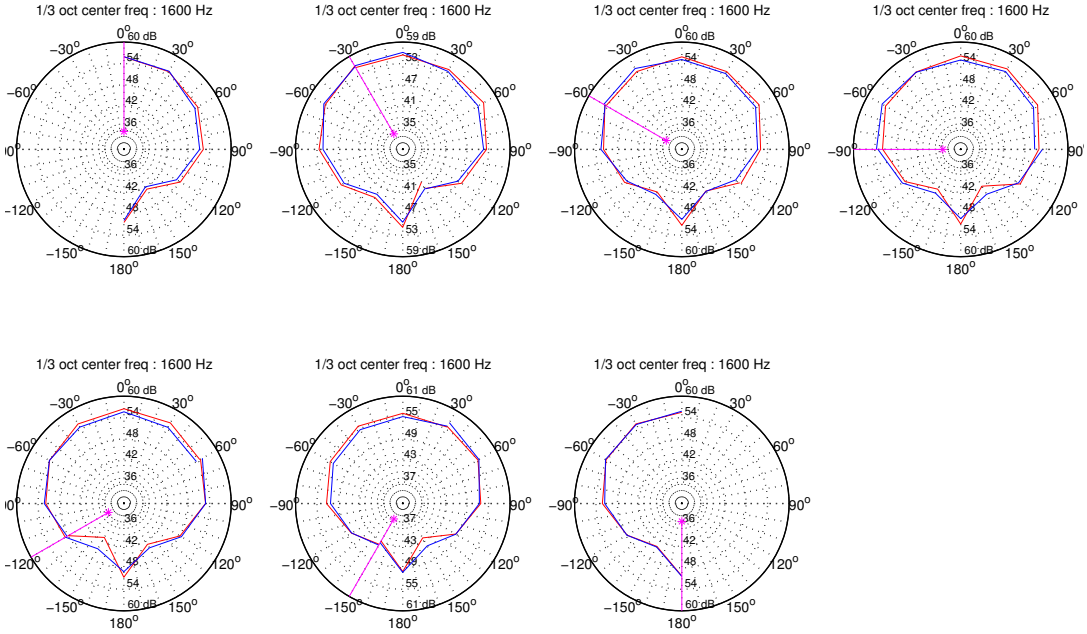


Figure A.13: Directivity comparison, 1600 Hz band.

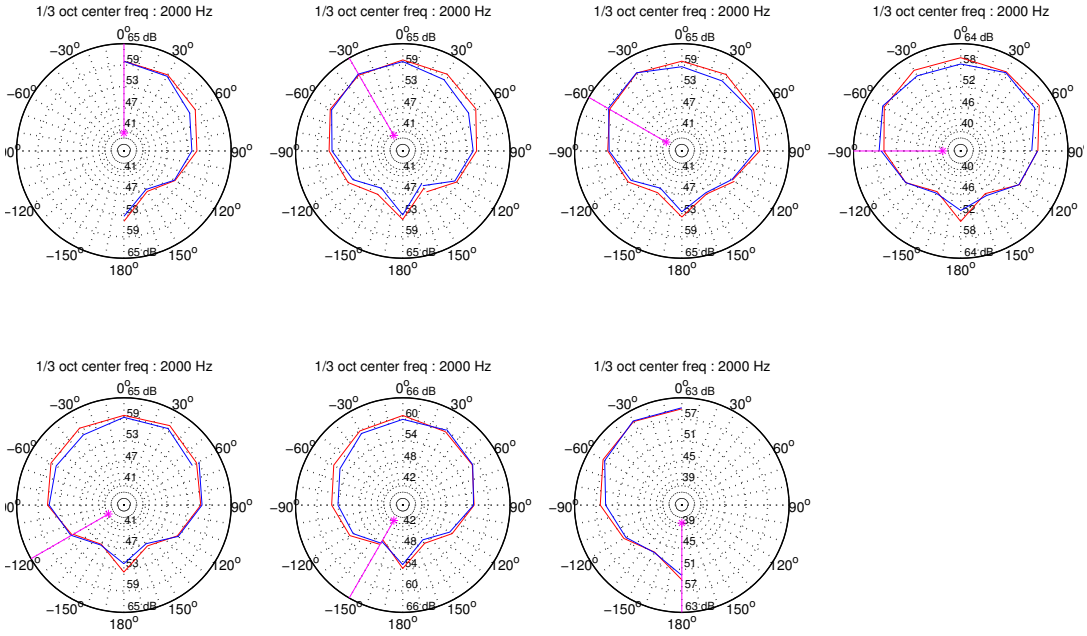


Figure A.14: Directivity comparison, 2000 Hz band.

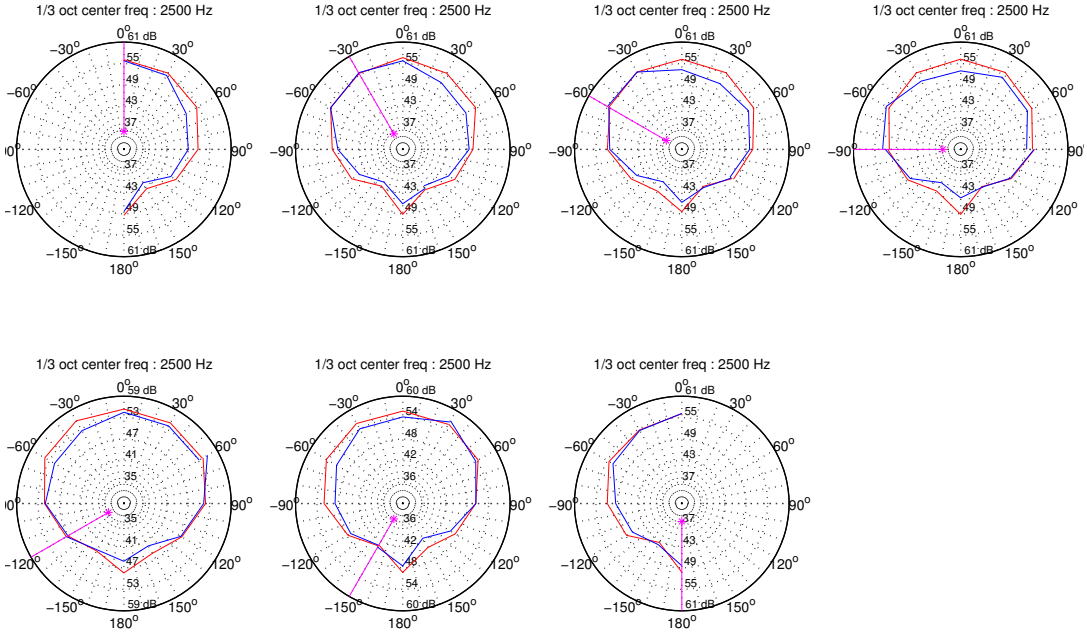


Figure A.15: Directivity comparison, 2500 Hz band.

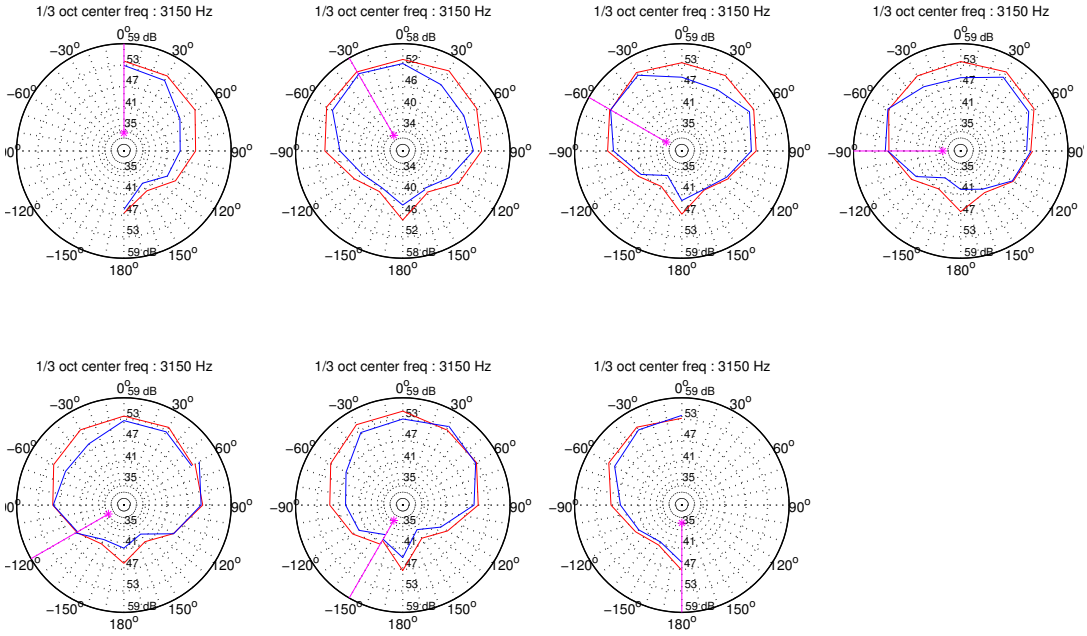


Figure A.16: Directivity comparison, 3150 Hz band.

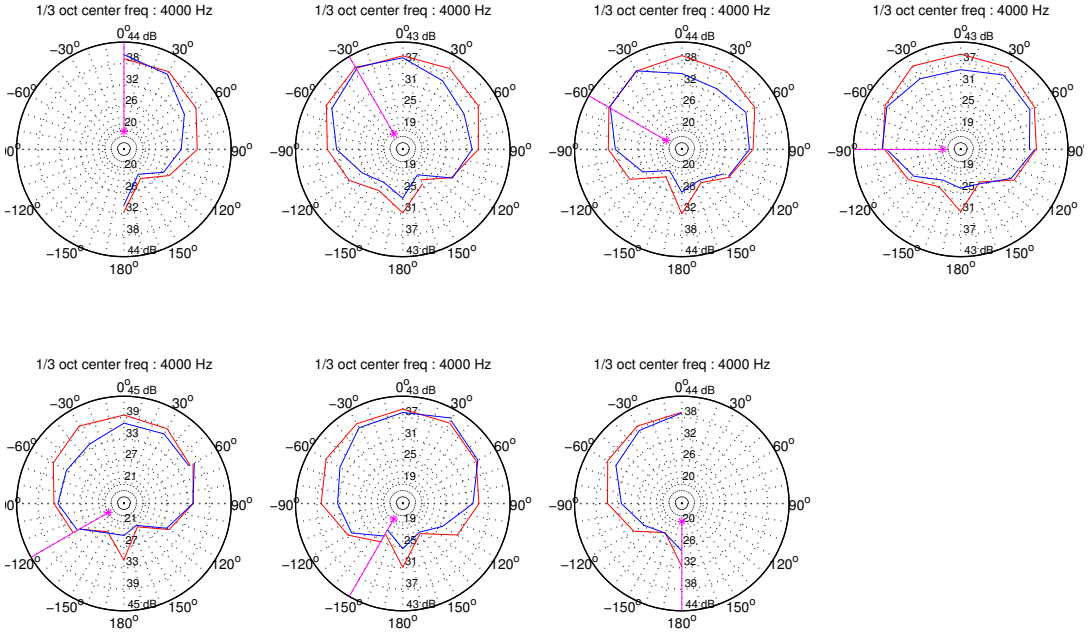


Figure A.17: Directivity comparison, 4000 Hz band.

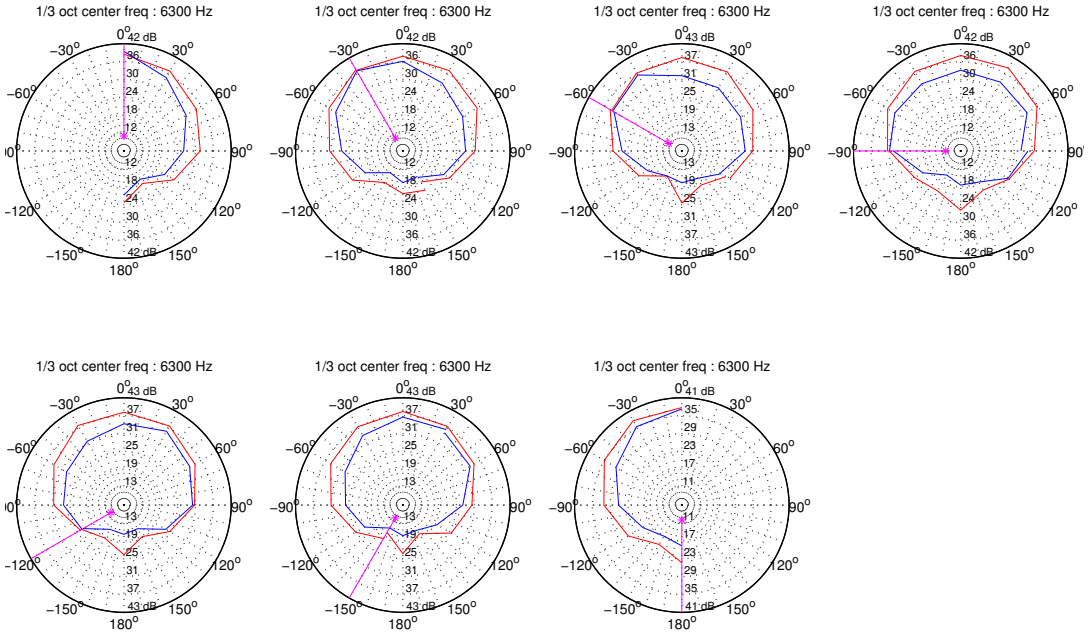


Figure A.18: Directivity comparison, 6300 Hz band.

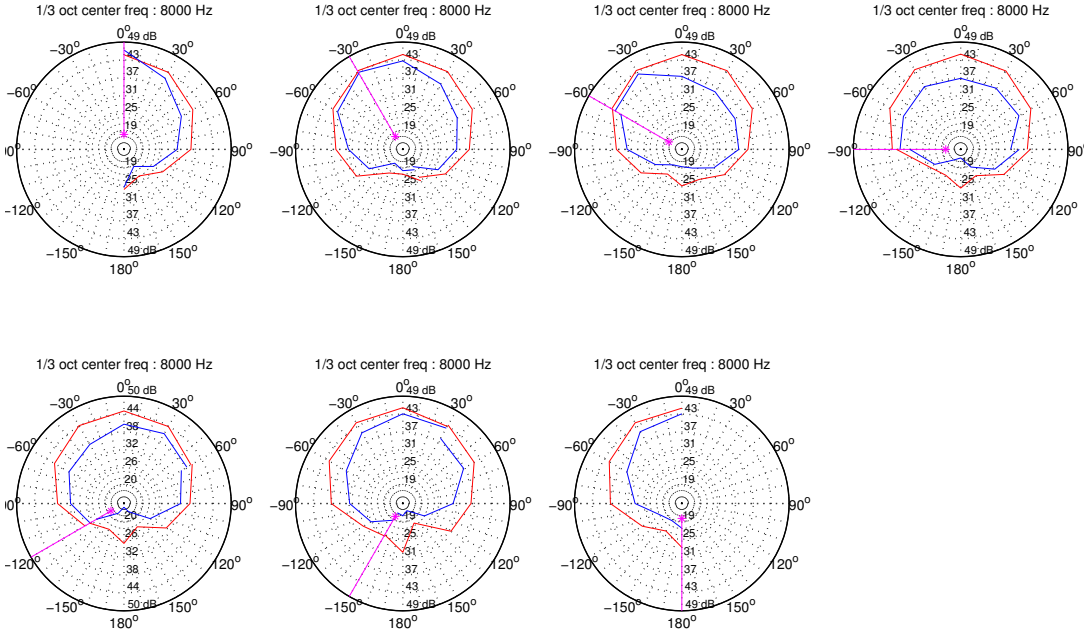


Figure A.19: Directivity comparison, 8000 Hz band.

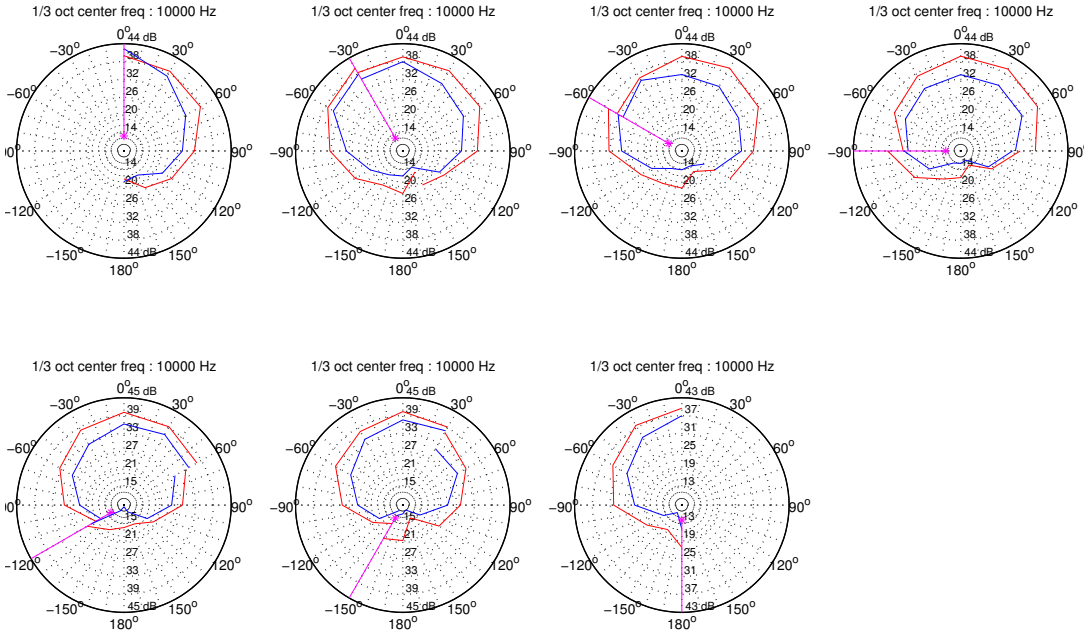


Figure A.20: Directivity comparison, 10000 Hz band.

A.2 Flow induced noise

This section presents more detailed measurement results of the binaural noise in individual wind direction. Figures A.22 to A.47 shows sound pressure level spectrum at the entrance of ear canals. One third octave frequency smoothing is applied to all plots. Figures 6.21 and 6.22 are based on this data. This section shows that there are some variations in spectrums from measurement to another within the same direction. Overall spectrum shape still remains the same.

The noise spectrum of the Knowles FG 23329 microphone used in the measurements are shown in figure A.21. Noise spectrum is measured in the big anechoic chamber of Department of Signal Processing and Acoustics where all the measurement were carried out.

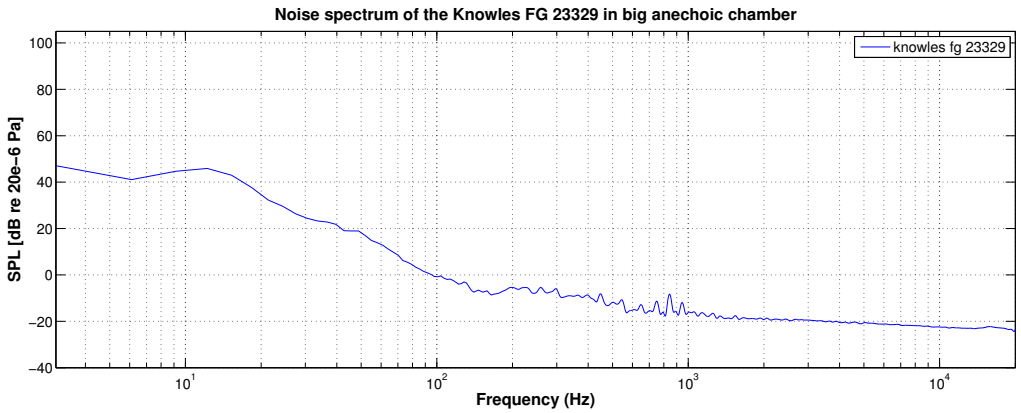


Figure A.21: Noise spectrum of the Knowles FG 23329 miniature microphone inside a big anechoic chamber.

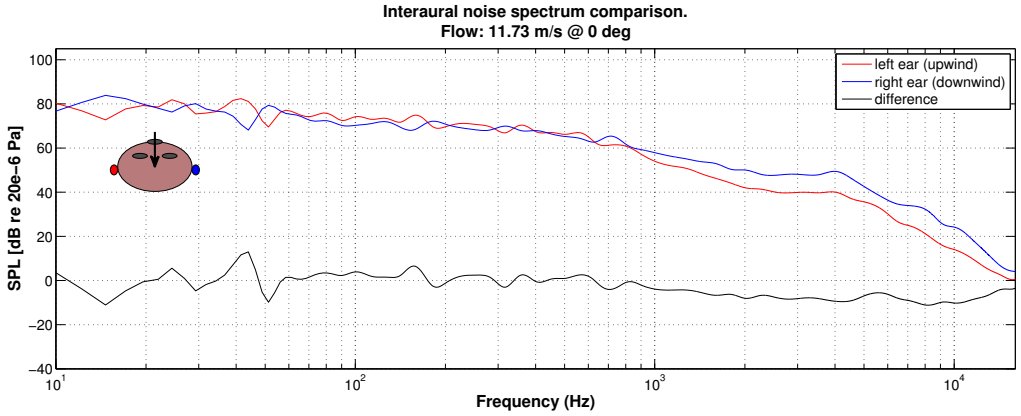


Figure A.22: Noise spectrum at the entrance of ear canals, 0°

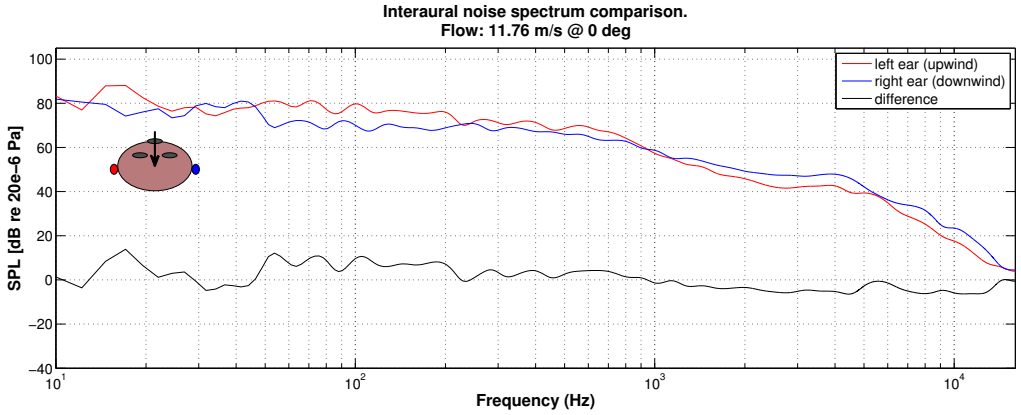


Figure A.23: Noise spectrum at the entrance of ear canals, 0°

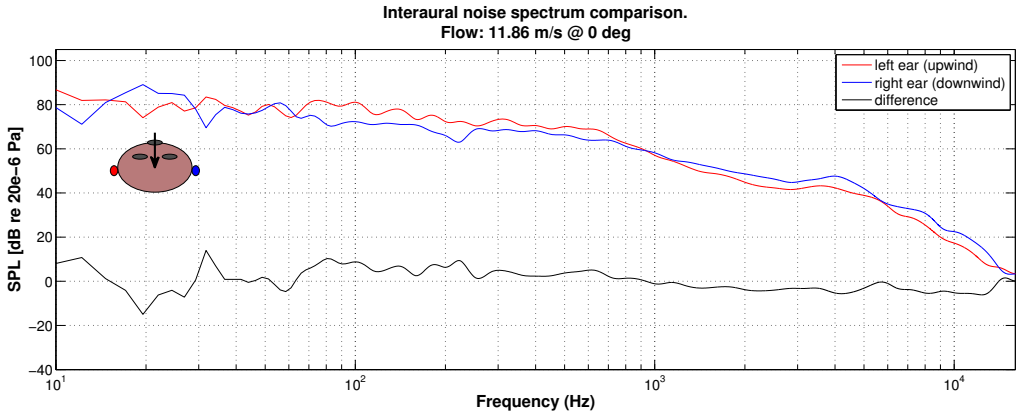


Figure A.24: Noise spectrum at the entrance of ear canals, 0°

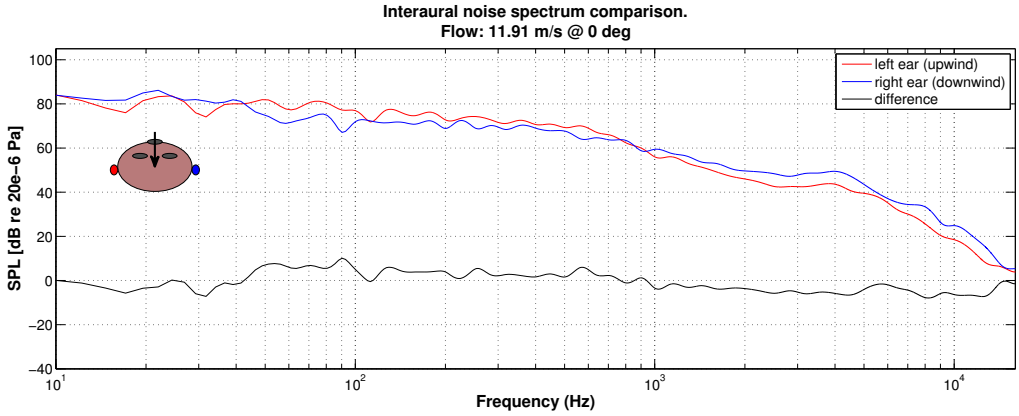


Figure A.25: Noise spectrum at the entrance of ear canals, 0°

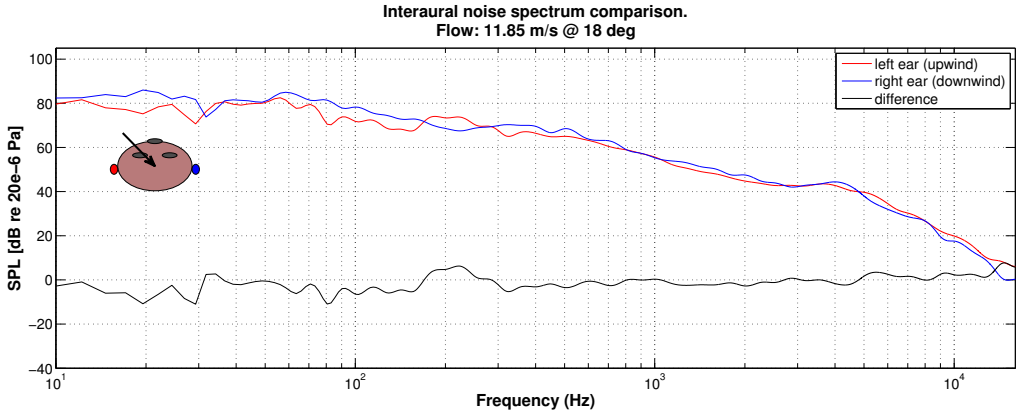


Figure A.26: Noise spectrum at the entrance of ear canals, 18°

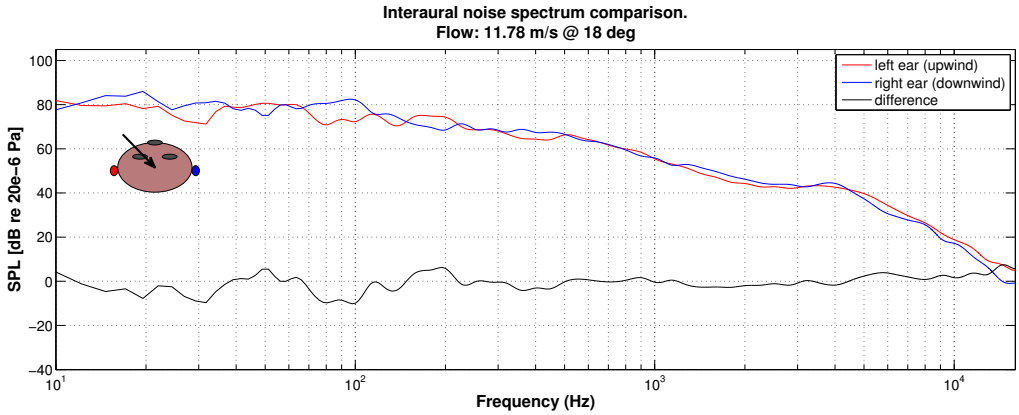


Figure A.27: Noise spectrum at the entrance of ear canals, 18°

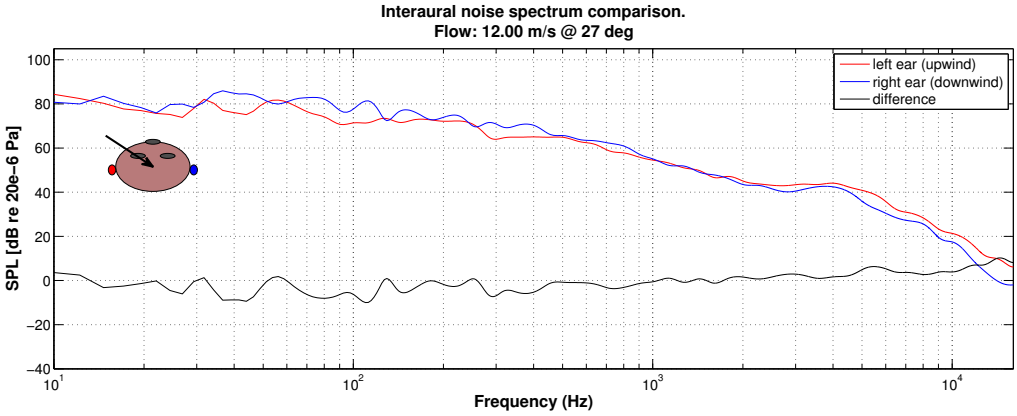


Figure A.28: Noise spectrum at the entrance of ear canals, 27°

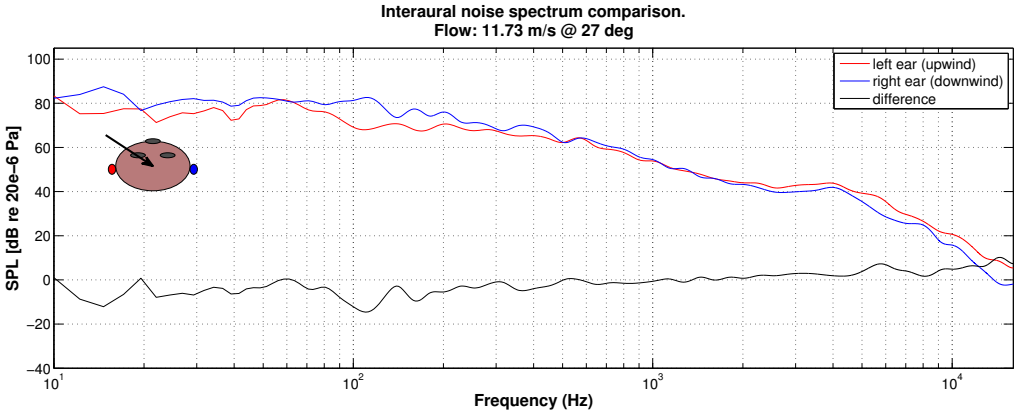


Figure A.29: Noise spectrum at the entrance of ear canals, 27°

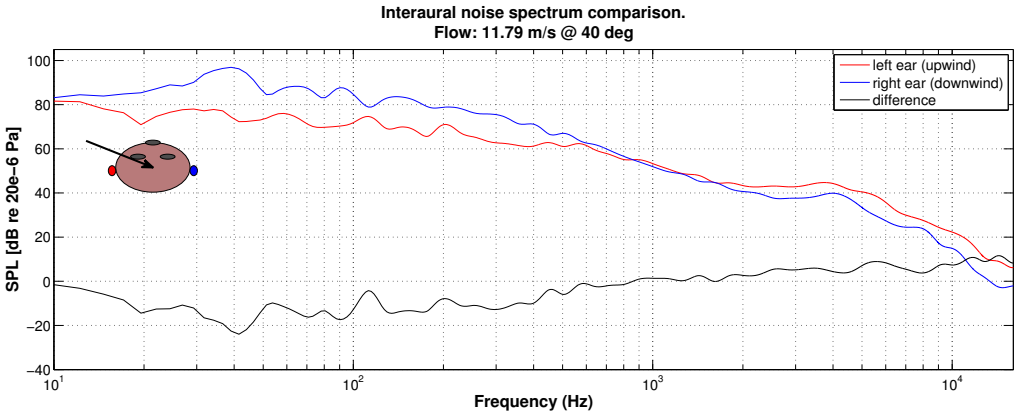


Figure A.30: Noise spectrum at the entrance of ear canals, 40°

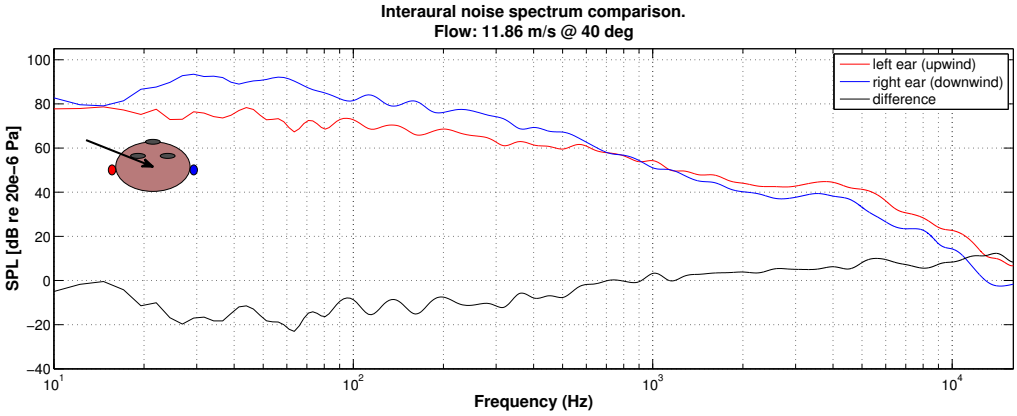


Figure A.31: Noise spectrum at the entrance of ear canals, 40°

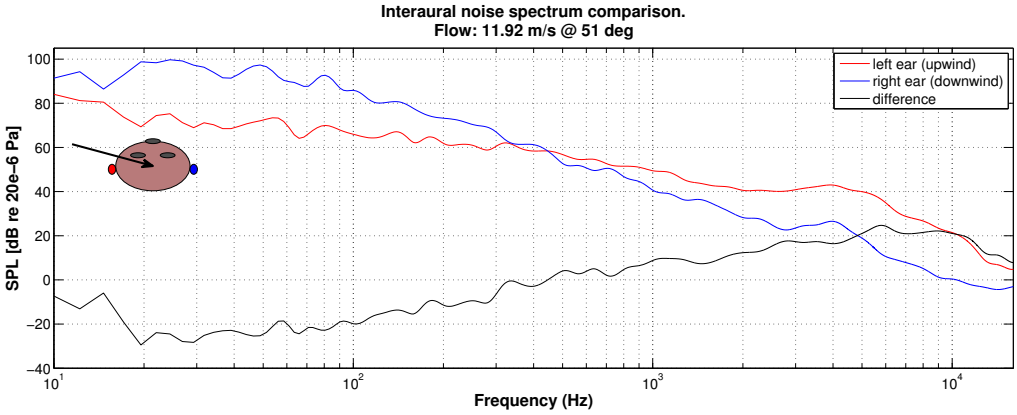


Figure A.32: Noise spectrum at the entrance of ear canals, 51°

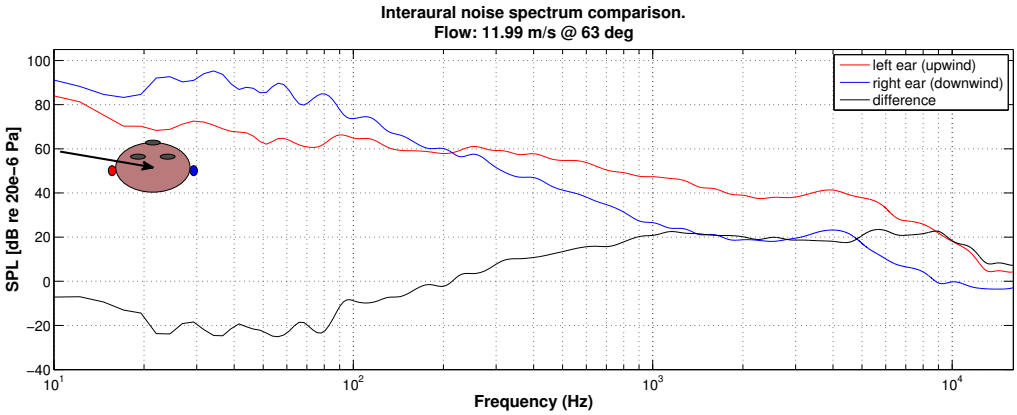


Figure A.33: Noise spectrum at the entrance of ear canals, 63°

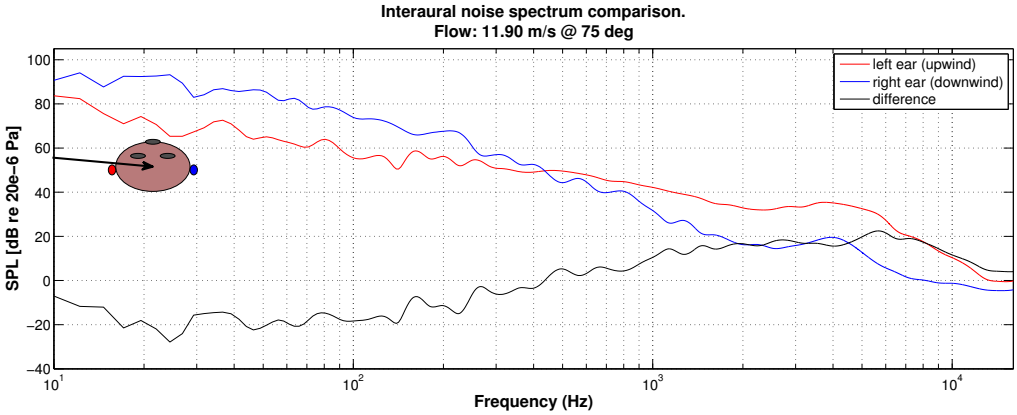


Figure A.34: Noise spectrum at the entrance of ear canals, 75°

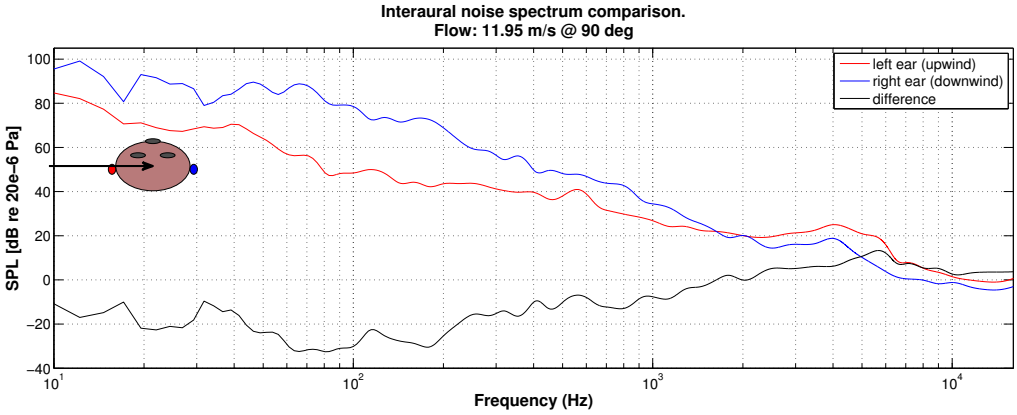


Figure A.35: Noise spectrum at the entrance of ear canals, 90°

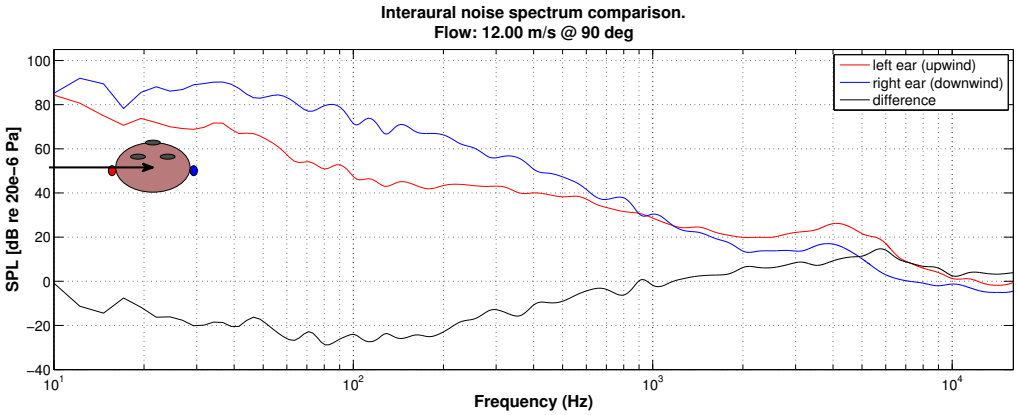


Figure A.36: Noise spectrum at the entrance of ear canals, 90°

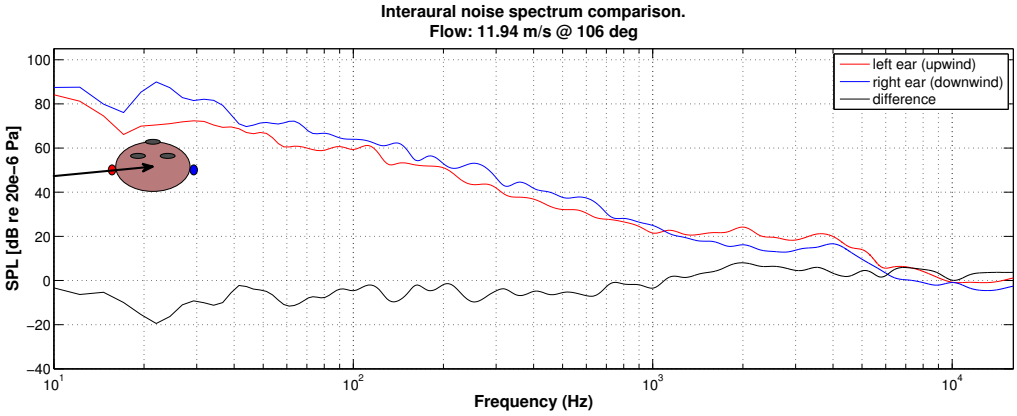


Figure A.37: Noise spectrum at the entrance of ear canals, 106°

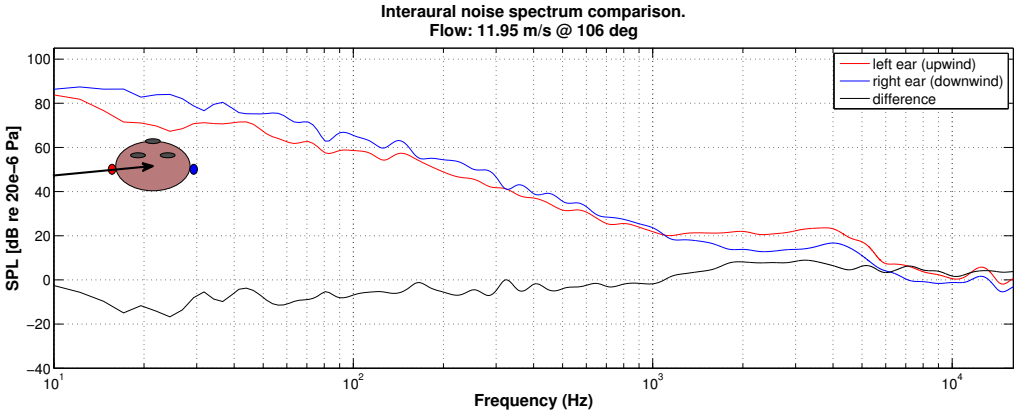


Figure A.38: Noise spectrum at the entrance of ear canals, 106°

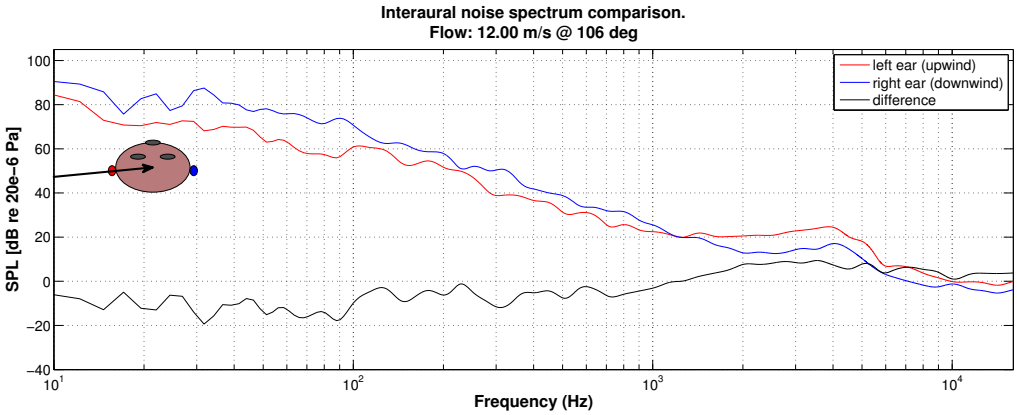


Figure A.39: Noise spectrum at the entrance of ear canals, 106°

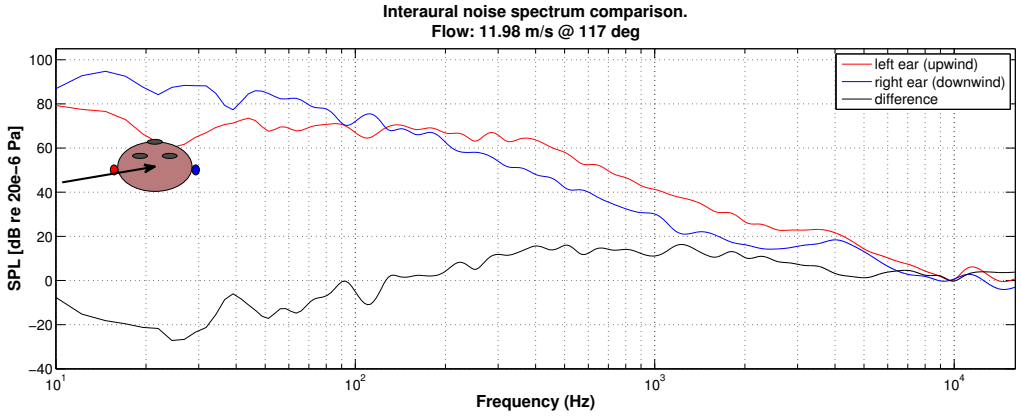


Figure A.40: Noise spectrum at the entrance of ear canals, 117°

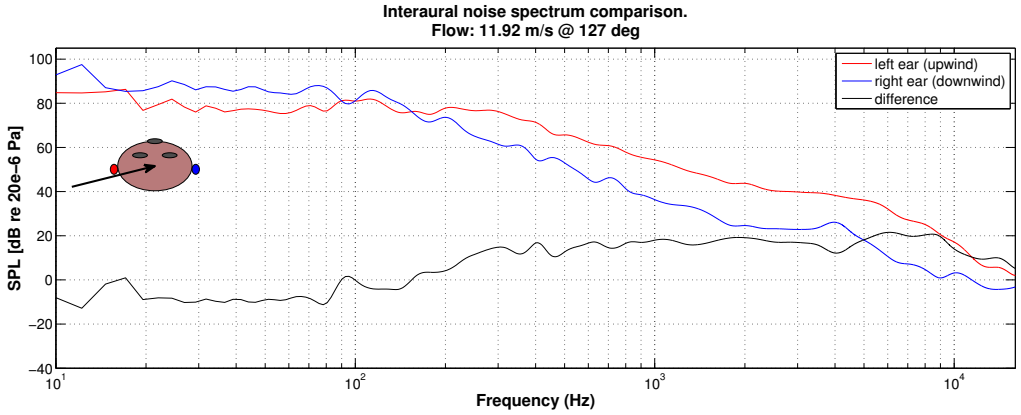


Figure A.41: Noise spectrum at the entrance of ear canals, 127°

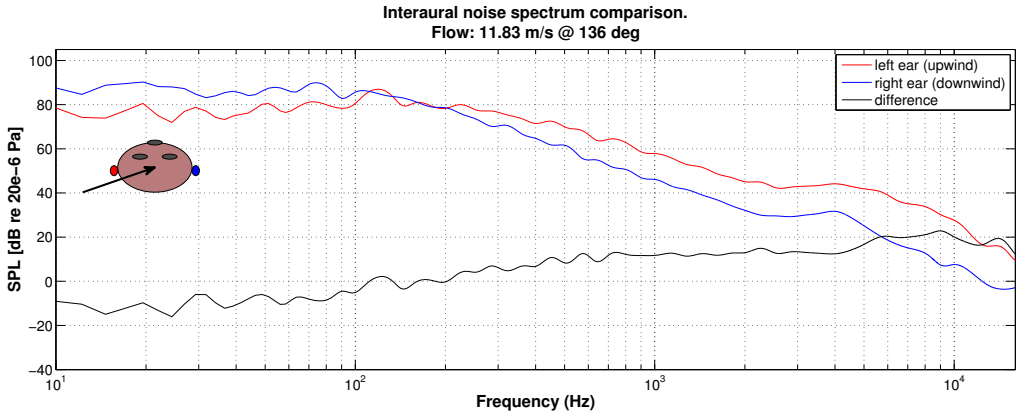


Figure A.42: Noise spectrum at the entrance of ear canals, 136°

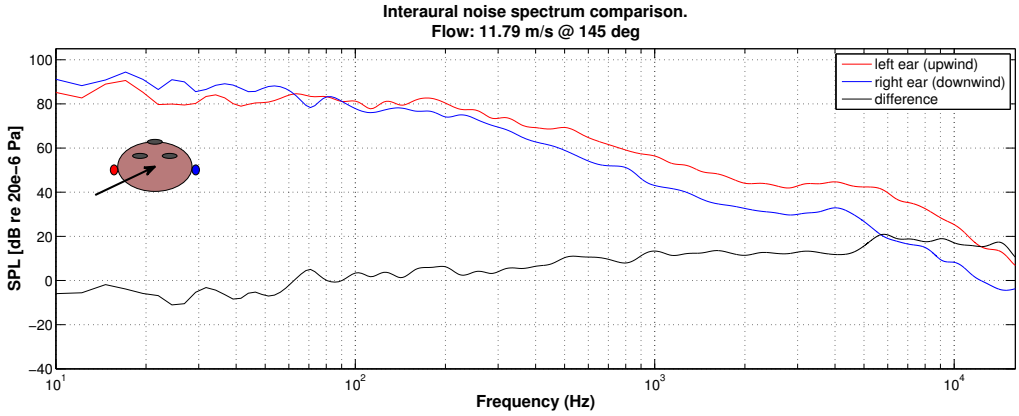


Figure A.43: Noise spectrum at the entrance of ear canals, 145°

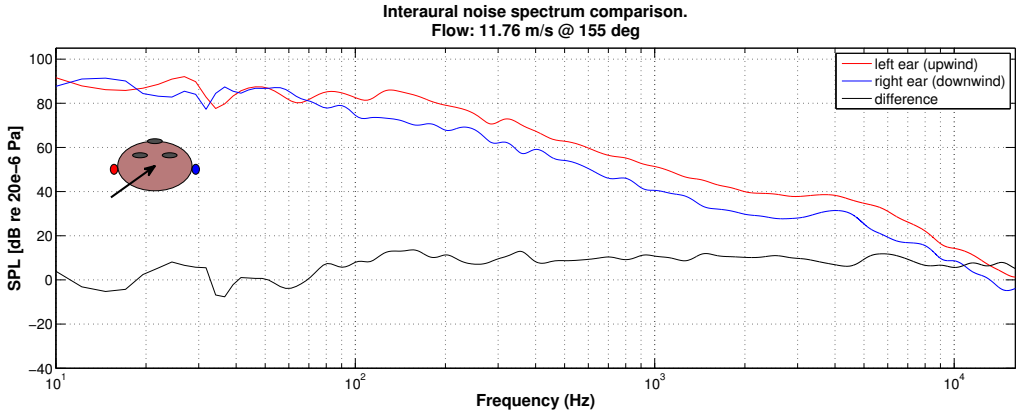


Figure A.44: Noise spectrum at the entrance of ear canals, 155°

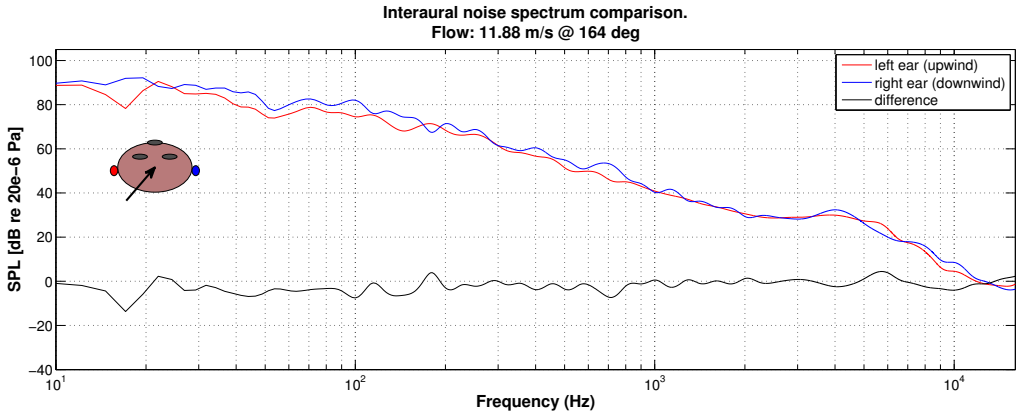


Figure A.45: Noise spectrum at the entrance of ear canals, 164°

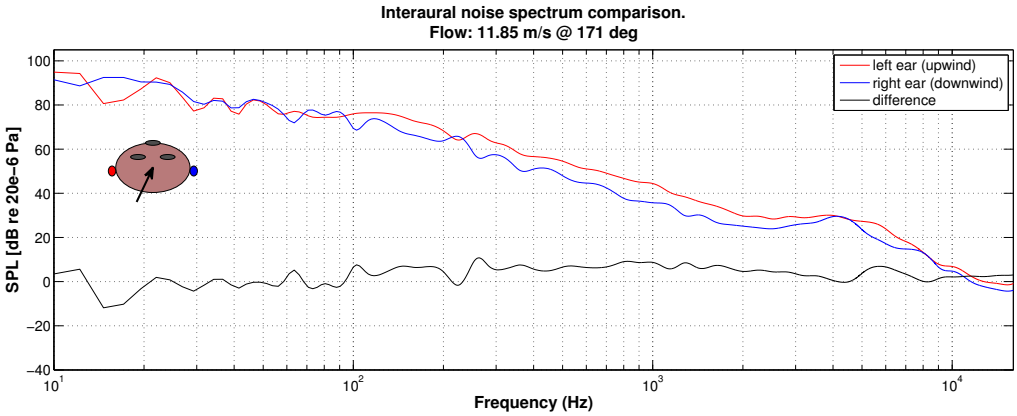


Figure A.46: Noise spectrum at the entrance of ear canals, 171°

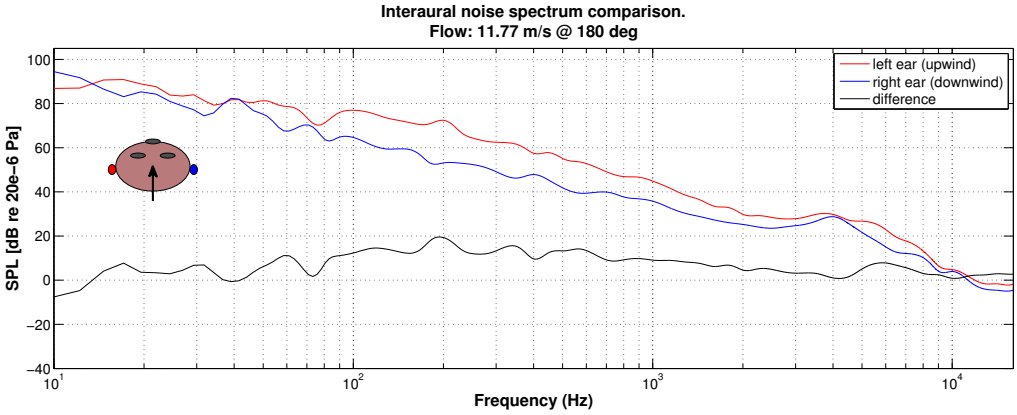


Figure A.47: Noise spectrum at the entrance of ear canals, 180°